This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Lexical and syntactic representations in closely related languages: Evidence from Cantonese–Mandarin bilinguals

Zhenguang G. Cai a,*, Martin J. Pickering a, Hao Yan b, Holly P. Branigan a

a Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, United Kingdom
b Faculty of English Language and Culture, Guangdong University of Foreign Studies, 2 Baiyun Dadao N., Guangzhou 510420, China

ABSTRACT
Bilinguals appear to have shared syntactic representations for similar constructions between languages but retain distinct representations for noncognate translation-equivalents (Schoonbaert, Hartsuiker, & Pickering, 2007). We inquire whether bilinguals have more integrated representations of cognate translation-equivalents. To investigate this, we report two structural priming experiments in which participants heard dative sentences in Mandarin or Cantonese and described pictures using Mandarin (Experiment 1) or Cantonese (Experiment 2). We found that cognate verbs between the prime and the target led to a smaller boost than same verbs. This difference in priming could not be attributed to more phonological overlap between same verbs (i.e., identical) than between cognate verbs (i.e., similar but not identical). These results suggest that cognate verbs have separate rather than shared lemma representations across languages, even though their associated syntactic information appears to be collectively represented. Furthermore, we found an advantage for within-language priming over between-language priming. We interpreted this advantage as the result of a language node passing activation to all the lemmas linked to it. Implications for bilingual lexical and syntactic representation and processing are discussed.

© 2011 Elsevier Inc. All rights reserved.

Introduction
To what extent do speakers of more than one language share lexical and syntactic representations between their languages? Previous work on this question has studied languages that are not very closely related, and suggests that certain syntactic representations are shared between the languages though lexical representations are maintained separately. In this paper, we ask whether speakers of two very closely related languages, Mandarin and Cantonese, make use of more fully integrated lexical and syntactic representations. To do this, we examine the occurrence of structural priming effects between Mandarin and Cantonese sentences that contain verbs that have the same or different meanings.

Most theories of bilingualism assume that bilinguals do not have entirely separate representations of their languages, but rather have a representation that is at least partly integrated (see Hartsuiker & Pickering, 2008). However, they tend to assume that this integrated representation draws on language-specific lexical representations; that is, words in a bilingual's lexicon are “tagged” for their language. In general, this appears necessary to prevent bilinguals regularly confusing their languages. But is such a separation always necessary? If two languages use words with similar forms and meanings, it would in principle be possible to represent those words once (with differences between the languages being indicated). Such words would therefore be shared between the languages. In linguistics, such words tend to be known as cognates, but the term can have rather different meanings to different researchers.
orthographically identical (e.g., “give”). In addition, all these translation-equivalents are bilinguals use separate production systems to draw on this (e.g., Kroll & Stewart, 1994). De Bot (1992) proposed that their two vocabularies in a single interconnected lexicon Lexical and syntactic representations in bilinguals have many characteristics that are usually associated with distinct languages, with their spoken forms being mutually unintelligible (Tang & van Heuven, 2009). But in many other respects they are closely related. Notably, they share a large part of the grammar and the same orthography (Mandarin and Cantonese generally use the same characters to express the same meanings). In fact, they are officially referred to by the Chinese Academy of Social Sciences as two dialects of the same language, Chinese (Xing, 1991).

Importantly, most Mandarin–Cantonese translation-equivalents are cognates in the sense that they are equivalent in meaning, similar in phonology (often in systematic and predictable ways), and identical in orthography. For instance, there is a systematic correspondence between the Mandarin consonant /hu/ (pronounced [xw] in IPA) and the Cantonese /w/ (pronounced [w] in IPA), as in huai ("bad"). Other words may differ only in the vowel (e.g., Mandarin di/Cantonese dai ("pass") or their tone (Mandarin song/Cantonese song ("give"). In addition, all these translation-equivalents are orthographically identical (e.g., di is used for both di and dai).

**Lexical and syntactic representations in bilinguals**

Substantial evidence suggests that bilinguals represent their two vocabularies in a single interconnected lexicon (e.g., Kroll & Stewart, 1994). De Bot (1992) proposed that bilinguals use separate production systems to draw on this single lexicon. However, these systems may interact with one another to an extent that depends on how closely the two languages are related, with closely related languages having closely integrated systems. Hartsook, Pickering, and Veltkamp (2004) proposed that bilinguals share syntactic representations between languages when the relevant structures are sufficiently similar in the two languages. Thus, the more similar that their syntax is, the more they share syntactic representations and processes. For example, the (surface) representations of the double-object (DO) and prepositional-object (PO) dative constructions are similar in Dutch and English: but the two languages differ substantially in other respects, such as word order in subordinate clauses (English: verb-medial; Dutch: verb-final). Bilingual speakers of these languages would therefore share syntactic representations for datives but not for verb position in subordinate clauses. Languages with more similar grammars would share proportionally more of their syntactic representations (see also Berneot, Hartsooker, & Pickering, 2007).

Schoonbaert et al. (2007) proposed one such model for Dutch–English bilinguals (see Fig. 1 for an adapted version in which language proficiency is ignored). This was based on Pickering and Branigan’s (1998) model of syntactic representation in monolinguals, which in turn drew on Levelt, Roelofs, and Meyer’s (1999) model of lexical representation. The model focuses on the lemma stratum, with each lemma representing the base form of a lexical item and its associated morphosyntactic information (Levelt et al., 1999). In the model, lemmas of the two languages are connected to the same conceptual representations but separate language nodes (see also Hartsooker & Pickering, 2008; Hartsooker et al., 2004). Lemmas are also linked to combinatorial nodes, representing the syntactic constructions in which they can occur. These syntactic representations, if sufficiently similar, are shared between languages (as the DO and PO constructions are here). In the model, translation-equivalent words have separate lemma representations but have shared representations for similar syntactic properties.

Schoonbaert et al. (2007) based their model on translation-equivalent verbs as a whole, and did not distinguish cognate translation-equivalents from noncognate ones (noncognates henceforth). But the distinction between cognates and noncognates might plausibly affect the nature

![Fig. 1. A model of lexico-syntactic representations adapted from Schoonbaert et al. (2007) for Dutch–English bilinguals.](image-url)
of lexico-syntactic representations in bilinguals. It has long been observed that cognates enjoy a closer relationship than noncognates in at least some aspects of bilingual language processing (for reviews, see Costa, Santesteban, & Cañó, 2005; Sanchez-Casas & Garcia-Albea, 2005). In comprehension, lexical access is facilitated for cognates but not (or to a lesser degree) for noncognates (cf. de Groot & Nas, 1991). For instance, Davis et al. (2010) found significant priming between English and Spanish cognates (e.g., rich-rico), which was not dependent on the degree of form overlap (i.e., phonological/orthographic similarity), but no priming between noncognates (e.g., duck-pato). Moreover, Spanish-Catalan bilinguals showed comparable priming for a target word (e.g., coche; Spanish “car”) following cognate primes (cotxe; Catalan “car”) as following identity primes (Garcia-Albea, Sanchez-Casas, & Valero, 1996). In production, Catalan-Spanish bilinguals were faster naming pictures with cognate names in Spanish and Catalan (e.g., gato-gat; “cat”) than pictures with noncognate names (e.g., meso-taula; “table”), both in Spanish and Catalan (Costa, Caramazza, & Sebastián-Gallés, 2000).

Costa et al. (2000) suggested that the cognate facilitation effect arises at the phonological level (see also Costa et al., 2005). For instance, when a Catalan-Spanish bilingual names a cat in Spanish (i.e., gato), he also activates the non-selected word (i.e., gat) and that activation cascades to its phonological features (e.g., /g/ /a/ and /t/). For cognates, activation then spreads to the phonological features of the target word (e.g., /g/ /a/ and /t/), but this does not happen for noncognates. However, this account has difficulty explaining the finding that the cognate facilitation effect was unaffected by degree of form overlap (Davis et al., 2010).

Alternatively, cognates but not noncognates may sometimes have a shared lemma representation (see Sanchez-Casas & Garcia-Albea, 2005, for such a possibility). Assuming that lexical access in bilinguals is language non-selective (e.g., Thierry & Wu, 2007), the access of, say, a Spanish word also entails the activation of its Catalan translation-equivalent. If cognates but not noncognates have the same lemma representation, a cognate’s lemma representation would receive additional activation (from its cognate counterpart) relative to noncognates, yielding the facilitation effect observed by Costa et al. (2000). In fact, a recent ERP study by Strijkers, Costa, and Thierry (2010) found that the cognate facilitation effect arises relatively early in picture naming (i.e., 190 ms after picture onset). As lemma retrieval in picture naming is estimated to occur at around 150–275 ms (Indefrey & Levelt, 2004), this early ERP effect appears to suggest that the facilitation arises during lemma retrieval, and hence that cognates may have shared lemmas. Furthermore, as the shared lemma representation is independent of wordform, the cognate facilitation effect should not be dependent on the degree of form overlap (e.g., Davis et al., 2010). Of course, it is also possible that the cognate facilitation effect arises from a shared morphological stem instead of (or in addition to) a shared lemma (Kirsner, Lalor, & Hird, 1993); however, in this paper we focus on the lemma representation.

Existing evidence regarding cognates is mainly concerned with wordforms, and does not indicate whether cognates have shared or separate lemma representations. Schoonbaert et al.’s (2007) model (see Fig. 1) assumes separate lemma representations for all translation-equivalents. However, the model was based on evidence from English and Dutch, two clearly distinct languages that have a relatively limited number of cognates; in fact, of the 6 pairs of Dutch–English verbs that Schoonbaert et al. used, only one pair was cognate under our definition (i.e., related in both meaning and phonology: geef–give). Thus, though their model captures the evidence regarding noncognates, it remains an open question whether it also holds for cognates.

Representation of Mandarin–Cantonese cognates

We can contrast two accounts of cognate representations in Cantonese–Mandarin bilinguals. The separate lemma account (Fig. 2a) assumes that such bilinguals have different lemmas for cognates, just as for noncognates, as assumed in Schoonbaert et al. (2007). Alternatively, the shared lemma account (Fig. 2b) assumes that such bilinguals (and bilinguals of other very closely related pairs of languages) have shared lemmas for cognates wherever possible. As we have noted, Hartsuiker et al. (2004) proposed on the basis of cross-linguistic structural priming effects that bilinguals share combinatorial nodes between languages for those combinatorial possibilities that both languages allow (but have separate combinatorial nodes for combinatorial possibilities that differ between languages). In the same way, bilinguals might share lemmas between languages for cognate words that have similar lexico-syntactic properties, as is usually the case for Mandarin and Cantonese verbs (but would have separate lemmas for any cognates that do not have similar lexico-syntactic properties). Note that in both the separate and shared lemma accounts, lemmas are linked to the relevant conceptual representations, language nodes, and combinatorial nodes. The crucial difference is that cognates are always represented as different lemmas in one account but may sometimes (and in the case of very closely related languages such as Mandarin and Cantonese, often) be represented as the same lemma in the other account.

The two accounts therefore explain in different ways how speakers select wordforms from the appropriate language. In the separate lemma account, a speaker selects either a Mandarin or a Cantonese lemma, and this leads to activation of the appropriate wordform. Language selection in this account therefore occurs at the lemma level. In the shared lemma account, in contrast, the lemma is shared between Mandarin and Cantonese (for most verbs), but we assume that this shared lemma (e.g., /di/ and /dai/) is linked to two separate wordform nodes (/di/ and /dai/), which are also connected to their respective language nodes (i.e., /di/ to Mandarin and /dai/ to Cantonese). Thus a bilingual who chooses to use Mandarin activates the /di/ lemma and the Mandarin node, which in turn collectively activate the /di/ wordform node. Language selection in this account therefore occurs at the wordform level.

The two accounts also explain in different ways how speakers encode different structural preferences for each language, as when a particular construction has different
frequencies in the two languages. For example, the DO construction is much less frequent in Cantonese than in Mandarin (see Experiment 2 below). The separate lemma account can explain such differences in terms of the strength of the links between each verb lemma and the combinatorial nodes. For instance, if the Mandarin verb *di* occurs more frequently in the DO construction than its Cantonese cognate *dai*, such a difference can be explained by a stronger link between *di* and the DO node than between *dai* and the DO node. In the shared lemma account, there is only one lemma for *di* and *dai* (i.e., *di/dai* in Fig. 2b), and hence only one link to the DO node. However, the shared lemma account can capture frequency differences by assuming links between the language nodes and the combinatorial nodes. In this case, there would be a stronger link between the Mandarin node and the DO node than between the Cantonese node and the DO node. In this account, choice of a construction would therefore result from activation from the lemma node and the relevant language node.

Previous research on bilingual language use might appear to favor the shared lemma account over the separate lemma account. First, there are both theoretical and empirical arguments that bilingual speakers aim for cognitive economy with respect to representations wherever possible. Thus de Bot (1992) argued that the more closely related two languages are, the more speakers will share representations and processes, and Hartsuiker et al. (2004) provided evidence that bilinguals share syntactic representations when languages permit it. Under the same logic, we might therefore expect lemma representations to be shared whenever possible. This is particularly the case for Mandarin and Cantonese, where cognates tend to have very similar morphosyntactic properties. Take dative verbs for instance: They both allow the PO/DO alternation, as well as topicalization, passivization, the *ba*-construction, and missing arguments, and they can both be combined with similar aspectual particles to express the timing of action. Of

---

1 It has been suggested that the DO construction is grammatical only with a limited set of verbs such as *bei* in Hong Kong Cantonese (e.g., Tang, 1993). However, the DO construction seems to be acceptable with all the dative verbs used in this study in Guangzhou Cantonese, as can be seen in participants’ production of DO constructions in Guangzhou Cantonese in Experiment 2.
course, there may be some across-the-board differences between languages – as we will see below – and there may additionally be some verbs with idiosyncratic properties that could not be represented via the same lemma. But lemma sharing is in principle possible for most Mandarin–Cantonese cognate verbs and specifically for the dative verbs that we consider here. In addition, the fact that Mandarin and Cantonese cognates have the same written form and closely related phonology would also favor the development of shared representations during learning: A child learning Cantonese and Mandarin simultaneously would usually find that a word in one language has an equivalent in the other language that has minor phonological differences and could be used similarly. In principle, such a situation could lead to him or her representing the two words as two variants of the same base form (i.e., lemma). Such a learning generalization in terms of phonological similarity is more likely given recent evidence that phonology can affect syntactic aspects of formulation (e.g., Santesteban, Pickering, & McLean, 2010).

In addition, some recent evidence supports the shared lemma account. As reviewed above, similarity in wordform does not appear to predict the size of the cognate facilitation effect in lexical access (Davis et al., 2010). Additionally, electrophysiological evidence suggests that the cognate facilitation effect arises early in lexical access, presumably during lemma selection (Strijkers et al., 2010). Of course, we are not suggesting that all cognates need have shared lemmas in the shared lemma account; some cognates may differ so much in lexico-syntactic properties that they obviously cannot have a shared lemma. But the shared lemma account may account for most Mandarin–Cantonese cognates (such as the dative verbs we examined in this study).

Structural priming and lexico-syntactic representations

We now consider how these two accounts can be contrasted using a structural priming paradigm (see Pickering & Ferreira, 2008, for a review). Structural priming refers to the tendency for people to repeat utterance forms to which they have been recently exposed. Bock (1986) found that English-speaking participants are more likely to describe a picture using a PO form such as the girl is handing a paint-brush to the boy after hearing and repeating another otherwise unrelated PO sentence such as the rock star sold some cocaine to the undercover agent than a DO sentence with the same meaning (the rock star sold the undercover agent some cocaine). Such priming does not depend upon lexical repetition (Bock, 1986; Bock, 1989), and is unaffected by the repetition of closed class elements such as prepositions or verbal morphology (Bock, 1989; Bock & Loebell, 1990; Pickering & Branigan, 1998), though it is enhanced when the lexical heads (e.g., the main verb) of prime and target utterances are the same (the lexical boost; Pickering & Branigan, 1998) or semantically related (the semantic boost; Cleland & Pickering, 2003).

The enhancement of structural priming with identical or related lexical heads can be captured by Pickering and Branigan’s (1998) model of lexical representation. On their account, producing an utterance with a particular structure activates the relevant lemma node and combinatorial node, as well as the link between them. In this account, structural priming is due to residual activation of the combinatorial node; the lexical boost to structural priming is due to residual activation of the lemma node (e.g., give), the combinatorial node (e.g., PO), and the link between them. The semantic boost occurs because each lemma node is linked to a conceptual node at the conceptual level, and nodes representing semantically related concepts are linked; activation during production or comprehension of a particular concept (e.g., GOAT) leads to some activation of related concepts (e.g., SHEEP), causing co-activation of the SHEEP lemma and the relevant combinatorial node, and hence enhanced priming. Evidence concerning feedback from the wordform levels to the lemma level is less clear: There is no evidence for a boost when heads have overlapping but distinct phonology (e.g., from the ship that is red to the sheep that is red in noun phrase priming; see Cleland & Pickering, 2003), but there is a boost when heads have identical phonology (i.e., are homophones; Santesteban et al., 2010).

Evidence of structural priming between languages suggests that bilingual speakers integrate the syntactic representation of their two languages to at least some degree. For example, Loebell and Bock (2003) found that German–English bilinguals tended to use a PO or DO form in German after using the structurally equivalent form in English, and vice versa. Similarly, Hartsuiker et al. (2004) found that participants were more likely to use an English passive after a Spanish passive than a Spanish active (see also Heydell & Murray, 2000; Meijer & Fox Tree, 2003). These findings have led to the proposal that bilinguals collectively represent similar syntactic constructions across languages so that using a construction in one language facilitates the use of a corresponding construction in the other language via the same syntactic representation (Hartsuiker et al., 2004; Schoonbaert et al., 2007; see also Fig. 1 above).

Schoonbaert et al. (2007) further found a between-language translation-equivalence boost: Although between-language priming occurred from Dutch to English and vice versa, it was stronger from Dutch (L1) to English (L2) for sentences that involved translation-equivalent verbs (which were mainly noncognates: e.g., gooien and throw) than non-translation-equivalent verbs (e.g., gooien and give). However, the translation-equivalence boost was only one seventh of the size of the within-language lexical boost. To account for these findings, Schoonbaert et al. proposed the separate lemma account for English and Dutch (noncognate) translation-equivalent verbs (as in Fig. 1). In this account, the translation-equivalence boost occurs because activation of the lemma of the verb in the prime sentence in a particular structure yields co-activation of the lemma of its translation-equivalent (via the shared conceptual node) and the relevant combinatorial node; this is similar to the mechanism underlying the semantic boost (e.g., between goat and sheep) proposed by Cleland and Pickering (2003). (Note that Schoonbaert et al. observed the translation-equivalence boost only from L1 [Dutch] to L2 [English]; we return to this in the General Discussion.)

The finding that verbs with distinct but related lemmas induce a boost that is smaller than the lexical boost can help us contrast the separate lemma (Fig. 2a) and the
shared lemma (Fig. 2b) accounts of cognate representations. Consider the Mandarin and Cantonese dative alternation (1a–d), which in both languages is similar to the English dative alternation, with the PO form involving the preposition gei/bei (to)\(^2\) and the inverse order of arguments from the DO form.\(^3,4\)

1. a. Niuzai di/huan-gei shuishou yitiao xiangjiao.
   (Mandarin DO)
   cowboy pass/return sailor a banana

2. b. Niuzai di/huan-le yitiao xiangjiao gei shuishou.
   (Mandarin PO)
   cowboy pass/return-ASP a banana to sailor

   (Cantonese DO)
   cowboy pass/return a banana

   (Cantonese PO)
   cowboy pass/return-ASP a banana to sailor

How might such sentences prime participants' choice of syntax when describing an event such as a policeman passing a soldier a hat, which can be described using a DO or PO form in both Mandarin and Cantonese? Both the separate and shared lemma accounts of course predict within-language priming (e.g., a greater likelihood of using a Mandarin DO following [1a] than [1b]) and a lexical boost (greater priming when prime and target have the same verb than different verbs). Both accounts also predict between-language priming (e.g., a greater likelihood of using a Mandarin DO following [1c] than [1d]), because under both accounts combinatorial nodes are shared between languages. Moreover, both accounts predict a cognate translation-equivalence boost (henceforth the cognate boost)\(^5\) to priming (i.e., greater between-language priming when the prime and target have cognate verbs than when they have different-meaning verbs), because on both accounts, a prime containing a particular verb in a particular structure will lead to co-activation of the lemma of its translation-equivalent (via the shared conceptual node) and the relevant combinatorial node, yielding enhanced priming.

However, the two accounts make different predictions concerning the magnitude of the cognate boost. Under the separate lemma account, the cognate boost should be smaller than the lexical boost (as observed by Schoonbaert et al., 2007). For example, when the speaker produces a Mandarin sentence using the verb di, the cognate boost from the daisentence compared to the wan-sentence in (1c) should be smaller than the lexical boost from the di-sentence compared to the huan-sentence in (1a). This is because when the prime sentence involves dai, the dai lemma becomes most activated and the separate di lemma is activated to a smaller extent (through the shared conceptual node). In contrast, under the shared lemma account the cognate boost should be comparable to the lexical boost, because the same lemma is activated by a prime containing dai as by a prime containing di.

Thus, we can examine whether cognates in Mandarin and Cantonese have shared or distinct lemma representations by comparing the extent to which priming is increased by repeating the same verb between sentences (i.e., the lexical boost) vs. the extent to which priming is increased by repeating a cognate verb with the same meaning between sentences (i.e., the cognate boost). In Experiment 1, participants produced Mandarin target descriptions after hearing Mandarin and Cantonese prime sentences; in Experiment 2, they produced Cantonese target descriptions after hearing Mandarin and Cantonese prime sentences. As well as manipulating prime structure (DO vs. PO) and priming mode (within- vs. between-language priming), we manipulated whether prime and target verbs had the same meaning (same verbs in between-language priming and cognate verbs in between-language priming) or different meanings. We predicted within- and between-language structural priming, together with a lexical boost and a cognate boost. But as we have seen, comparison of the magnitude of the cognate boost (from cognate verbs) as compared to that of the lexical boost (from same verbs) should allow us to distinguish the separate and shared lemma accounts.

**Experiment 1**

**Method**

**Participants**

Seventy-two college students in Guangzhou were paid 15 RMB (roughly 2 dollars) to take part. They were native speakers of Cantonese living in a predominantly Cantonese-speaking region but had used Mandarin since early childhood, and used both languages daily.

**Materials**

We created 32 experimental items (see Appendix A) and 96 filler items. Each item comprised a prime sentence, a prime picture, and a target picture (see Fig. 3). An experimental prime sentence had 8 versions: a Mandarin DO sentence, a Mandarin PO sentence, a Cantonese DO sentence, and a Cantonese PO sentence (see 1a–d); each of the sentences also had a same-meaning verb version (i.e., same-meaning verbs between prime and target) and
a different-meaning verb version (i.e., different-meaning verbs between prime and target). The filler prime sentences included 48 active transitives and 48 intransitives. Half of the transitives and intransitives were in Mandarin and the other half were in Cantonese. The prime sentences were read by a female speaker from the same population as the participants and were digitally recorded as WAV files in a recording studio.

The prime pictures depicted entities interacting; a verb indicating the action was printed below the picture. Half of the depicted events matched the event described by the corresponding prime sentence; the other half did not. For the mismatching experimental prime pictures, the Agent, the Theme, or the Recipient in the event was incompatible with the sentence; for the mismatching transitive filler prime pictures, either the Agent or the Patient was incompatible; for the mismatching intransitive filler prime pictures, the Agent was incompatible.

The target picture depicted an event that was unrelated to the event expressed in the prime sentence or prime picture and did not involve any of the same entities. It included a sentence preamble and an underline below the depicted event. For experimental target pictures and filler target pictures that depicted a transitive event, the preamble included the Agent and the verb. The use of the preamble in the experimental target pictures helped induce DO or PO utterances (by preventing the use of other alternative constructions such as the ba-construction). For filler target pictures that depicted an intransitive event, the preamble contained the subject only.

Procedure

Instructions were given in Mandarin. Participants were told to produce target completions in Mandarin. The experiment was run on a computer using DMDX (Forster & Forster, 2003). Items were presented in individually randomized orders so that 2–4 fillers separated experimental items. Participants were first shown the entities (e.g., a pirate, a ball) that were to appear in the experiment, together with their names, in PowerPoint slides. After this, participants were presented with 3 practice items. For each item (in both the practice and the real experiment), participants first saw a line of dashes on the computer screen. After participants pressed the spacebar, the prime sentence was played. The prime picture then appeared and participants decided whether it matched the prime sentence by pressing F (match) or J (mismatch). The target picture then appeared and participants described it using the preamble provided as the beginning of their description (see Fig. 3). The experiment took about 45 min.

Scoring

Following previous studies (e.g., Pickering & Branigan, 1998), all responses were scored as DOs, POs, or Others. A response was encoded as DO when the verb was followed by the Recipient and then the Theme (e.g., jingcha di-gei shi-bing yiding maoizi, “the policeman passed the soldier a hat”), and as PO when the verb was followed by the Theme and then the Recipient (e.g., jingcha di-le yiding maozi gei shibing, “the policeman passed a hat to the soldier”). The rest of the responses were encoded as Others (which include cases where the preamble was changed, the response was ungrammatical, the Theme or the Recipient was omitted, or there was no response).

Data analysis

We carried out analyses that compared the number of primed responses (where the target had the same structure as the prime; i.e., a DO response following a DO prime or a PO response following a PO prime) and unprimed responses (where the target had the alternative structure to the prime; i.e., a DO response following a PO prime or a PO response following a DO prime). Thus, structural priming is measured by the relative frequencies of primed vs. unprimed responses, rather than as an interaction between
prime construction and target construction (e.g., Pickering & Branigan, 1998); in keeping with this approach, the lexical boost was captured as whether priming was greater following same-verb primes than following different-verb primes, rather than as an interaction of prime construction (PO vs. DO) and verb (same vs. different). Analyzing the data in this way reduced the complexity of the analyses. For convenience, we also report the priming effect for each condition, calculated as the primed responses in each condition divided by the sum of primed and unprimed responses in that condition.

In the statistical analysis, we used logit mixed effects (LME) modeling (e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008; Jaeger, 2008), and followed the analytic method adopted in Baayen (2008). We first applied centering to the fixed predictors, verb meaning and priming mode, assigning numeric values with a range of 1 and a mean of 0 to levels within a predictor. We incorporated all the fixed predictors and their interactions in a model in a single step. The model also included participants and items as random intercepts. Furthermore, we considered whether participants or items differed in their slopes for any fixed effect by using forward model comparison: They did if their inclusion significantly improved the model fit (Baayen et al., 2008; see also Sturt, Keller, & Dubey, 2010). In the statistical report, no random participant/item slopes were included in a model except for those that are mentioned. All the analyses were done in the software R (Version 2.10.1).

Results and discussion

Table 1 reports DO, PO, and Other responses in each condition. Table 2 reports the results of the analyses. The intercept indicates that there were significantly more primed responses than unprimed responses, indicating that participants tended to use the same structure as the prime when describing the target picture. In other words, structural priming occurred. Verb meaning had a main effect, with same-meaning verbs inducing greater priming than different-meaning verbs (.78 vs. .61). The effect held both when the prime and the target were in the same language and when they were not (see Fig. 4). There was a larger within-language priming effect with the same verbs than with different-meaning verbs (Estimate = 1.20, SE = .16, z = 7.29, p < .001, participants differed in slopes for verb meaning); that is, there was a lexical boost to priming. There was also a larger between-language priming effect with cognate verbs than with different-meaning verbs (Estimate = .66, SE = .13, z = 5.16, p < .001); that is, there was a cognate boost to priming (see Fig. 4).

There was also a main effect of priming mode, with greater within- than between-language priming (.74 vs. .65). For same-meaning verbs, the priming effect was greater when the prime and the target used the same verb than when they used cognate verbs (Estimate = .68, SE = .15, z = 4.48, p < .001). For different-meaning verbs, the priming

### Table 1

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Target responses and priming effect by prime condition in Experiment 1.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Same-meaning verbs</td>
</tr>
<tr>
<td></td>
<td>Within language</td>
</tr>
<tr>
<td>DO</td>
<td>195</td>
</tr>
<tr>
<td>PO</td>
<td>84</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
</tr>
<tr>
<td>Primed</td>
<td>466</td>
</tr>
<tr>
<td>Unprimed</td>
<td>96</td>
</tr>
<tr>
<td>Priming</td>
<td>.83</td>
</tr>
</tbody>
</table>
effect was greater when the prime and the target were in the same language than when they were in different languages (Estimate = .68, SE = .15, z = 4.48, p < .001); in other words, there was a within-language advantage for unrelated verbs.

There was a significant interaction between verb meaning and priming mode: Same-meaning verbs yielded a greater boost (relative to different-meaning verbs) when the prime and the target were in the same language than when they were in different languages (see Fig. 4). Hence there was a greater boost when the prime and target involved the same verb (i.e., the lexical boost) than when they involved cognate verbs (i.e., the cognate boost). (It should be noted that the difference between the lexical boost and the cognate boost was not an effect of the general advantage of within-over between-language priming: The lexical boost was the difference in priming between same-meaning verbs and different-meaning verbs in the within-language conditions.) Finally, analyses of Other responses showed no main effects of verb meaning or priming mode, nor any significant interaction.

These results are therefore incompatible with the shared lemma account, which does not predict a difference between the lexical boost and the cognate boost. Instead, they support the separate lemma account. Note that the finding of greater within-than between-language priming for different-meaning verbs contrasts with Schoonbaert et al.’s (2007) finding of comparable within-language and between-language priming for sentences involving different-meaning verbs (see also Kantola & Van Gompel, 2011); we will return to this issue in the General Discussion.

We note that there appeared to be some asymmetry in the priming effect following DO and PO primes: The boosts seem to have been mainly driven by DO primes rather than PO primes in this experiment (and also in Experiment 2 below). We attribute this asymmetry to a ceiling effect following PO primes, which induced almost entirely PO responses (over 86%; see also Experiment 2). Such an asymmetry has likewise been found in other syntactic priming studies when the two alternate structures differ substantially in their frequencies (e.g., Cleland & Pickering, 2003.)

Experiment 2 was identical to Experiment 1, except that Cantonese was the target language. Experiment 2 therefore provided a further test of the shared and separate lemma accounts, and additionally allowed us to compare the effects of Mandarin and Cantonese as target languages.

**Experiment 2**

**Method**

**Participants**

Seventy-two further participants from the same population as the participants in Experiment 1 were paid 15 RMB (roughly 2 dollars) to take part. Seven participants were replaced because they produced Other responses for more than one-third of targets or because of technical problems such as recording failures.

**Materials, procedure, and scoring**

These were the same as Experiment 1, except that participants were asked to describe the target picture in Cantonese. We gave the experimental instructions in Mandarin, in part for consistency with Experiment 1, and in part because Mandarin is normatively more appropriate in academic situations.

**Results and discussion**

Table 3 presents the descriptive results of the experiment and Table 4 presents results of the analyses. As shown in Table 4, the significant intercept indicates more primed than unprimed responses: Participants tended to use the same structure as the prime when describing the target picture. Same-meaning verbs induced greater priming than different-meaning verbs (.66 vs. .55), which is true for both within-language priming (Estimate = .61, SE = .13, z = 4.78, p < .001) and between-language priming (Estimate = .27, SE = .12, z = 2.21, p < .05) (see Fig. 5). These findings suggest both a lexical boost from same verbs and a cognate boost from cognate verbs. Priming mode also produced a main effect: There was greater within- than between-language priming (.63 vs. .59). Separate analyses show that the effect of priming mode held with same-meaning verbs (Estimate = .36, SE = .13, z = 2.8, p < .01), suggesting that the priming was greater when the prime and the target used the same verbs (i.e., within-language priming) than when they used cognate verbs (i.e., in-between-language priming). However, unlike in Experiment 1, the priming mode effect did not hold with different-meaning verbs (Estimate = .02, SE = .12, z = .12, p > .1), which suggests that there was no within-language advantage in priming for different-meaning verbs. There was a significant interaction in part because Mandarin is normatively more appropriate in academic situations.

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Same-meaning verbs</th>
<th>Different-meaning verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Within language</td>
<td>Between language</td>
</tr>
<tr>
<td></td>
<td>DO PO</td>
<td>DO PO</td>
</tr>
<tr>
<td>DO</td>
<td>110 3</td>
<td>69 5</td>
</tr>
<tr>
<td>PO</td>
<td>166 276</td>
<td>203 268</td>
</tr>
<tr>
<td>Other</td>
<td>12 9</td>
<td>16 15</td>
</tr>
<tr>
<td>Primed</td>
<td>386 337</td>
<td>294 292</td>
</tr>
<tr>
<td>Unprimed</td>
<td>169 208</td>
<td>235 237</td>
</tr>
<tr>
<td>Priming</td>
<td>.70 .62</td>
<td>.56 .55</td>
</tr>
</tbody>
</table>

### Table 4

LME results for Experiment 2.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>z</th>
<th>p</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.46</td>
<td>.06</td>
<td>7.34</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Verb meaning</td>
<td>.48</td>
<td>.10</td>
<td>4.85</td>
<td>&lt;.001</td>
<td>(p)</td>
</tr>
<tr>
<td>Priming mode</td>
<td>.19</td>
<td>.09</td>
<td>2.14</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Verb meaning × priming mode</td>
<td>.35</td>
<td>.18</td>
<td>1.94</td>
<td>.05</td>
<td></td>
</tr>
</tbody>
</table>

Note: The “Slope” column indicates whether a model included the random slope parameter corresponding to the effect for participants (p) or items (i), or it included no random slope (p).
between verb meaning and priming mode, which suggests that there was a greater boost from same-meaning verbs (relative to different-meaning verbs) in within-language priming than in between-language priming (see Fig. 5); in other words, the lexical boost was greater than the cognate boost.

As in Experiment 1, Experiment 2 showed a general structural priming effect, this time when Cantonese was the target language. The finding that cognate verbs induced a smaller boost than same verbs suggests that cognate verbs are represented as distinct lemmas, supporting the separate lemma account. Unlike Experiment 1, Experiment 2 did not show greater within- than between-language priming for different-meaning verbs. This null effect may reflect the fact that participants used the PO structure about 90% of the time for different-meaning verbs. Across both same- and different-meaning verbs, the strong tendency to produce PO responses may have underlain the relatively small priming effects (i.e., there was a ceiling effect).

We additionally carried out combined analyses over the pooled data from Experiments 1 and 2, using response (DO vs. PO) as the dependent variable and target language (Mandarin vs. Cantonese) as the fixed predictor, in order to examine whether the frequencies of DO and PO responses differed between Mandarin and Cantonese. There were more PO responses (.77) than DO responses (.23) across the two experiments (i.e., for Mandarin and Cantonese responses collapsed together). More importantly, target language significantly predicted the distribution of the responses (Estimate = -1.77, SE = .25, z = -6.97, p < .001, participants differed in slopes for target language), with more DOs in Mandarin (.33) than in Cantonese (.14), and more POs in Cantonese (.86) than in Mandarin (.67).

Our data also allowed us to investigate whether there were consistent relationships between structural preferences for cognate verbs in Mandarin and Cantonese. To do this, we correlated the DO/PO preference (i.e., the likelihood for a verb to be used in a DO or PO construction) for each verb in Mandarin (Experiment 1) and Cantonese (Experiment 2). The DO/PO preference was highly correlated between Mandarin and Cantonese cognates (r = .70, p < .01); if a Mandarin verb had a high probability of occurring in a DO or PO construction, so did its Cantonese cognate. Thus, roughly half of the variation in DO/PO preference between the languages could be attributed to an across-the-board difference (even though the two experiments involved different participants). Hence there was a systematic shift in DO/PO preference between the two languages.

**Phonological similarity rating study**

Same-meaning verbs in the between-language conditions often had different phonological forms (e.g., di and dai; “pass”). Thus it is conceivable that phonological similarity may underlie the stronger within- than between-language priming for same-meaning verbs, and by extension, the greater boost from same-meaning verbs (relative to different-meaning verbs) for within-language priming (i.e., the lexical boost) than for between-language priming (i.e., the cognate boost).

The evidence for a phonological boost to structural priming is mixed. Pickering and Branigan (1998) found that repeating the verb enhanced priming (the lexical boost) in sentence completion, but repeating the form of the verb (e.g., gave–gave vs. was giving–gave) did not enhance priming. This suggests that the lexical boost depends on lemma repetition rather than morphological or, importantly, phonological repetition. Similarly, Cleland and Pickering (2003) found no phonological boost to priming of noun–phrase structure (e.g., a ship–sheep prime-target pair did not enhance priming despite their phonological similarity). Finally, Bock and Loebell (1990) found no priming on the basis of metrical similarity alone (e.g., a sentence such as Susan brought a book to study did not prime PO responses). However, recent work suggests that there is a homophone boost, with participants being more likely to describe a flying bat as the bat that’s red after hearing the bat that’s red (where bat referred to a cricket bat) than after hearing the pool that’s red (Sanesteban et al., 2010). Full phonological overlap therefore appears to enhance structural priming, but this may be due to a shared wordform node (bat) rather than overlapping phonemes (as is the case in di vs. dai).

To test whether phonological similarity enhanced structural priming for same-meaning verbs, we asked 32 further participants from the same population as Experiments 1 and 2 to rate the phonological similarity of the cognate verb pairs (e.g., di - dai) used in the experiments on a scale from 1 (they sound very alike) to 7 (they sound very different); see Appendix B. We interleaved the 14 pairs of target cognate verbs with 14 filler pairs of Mandarin–Cantonese cognates, began the questionnaire with 8 practice items, and reversed the order of presentation for half the participants. To control for any ordering effects, we asked half the participants to compare the Mandarin pronunciation of the character against the Cantonese pronunciation and the other half to compare the Cantonese pronunciation of the character.
against the Mandarin pronunciation. As every verb is phonologically identical to itself, we did not ask participants to rate the phonological similarity of same verb pairs (e.g., \textit{di–di}); instead, we automatically assigned 1 to same-verb pairs.

Note that verb status (same verbs vs. cognate verbs) and phonological similarity should have high collinearity (i.e., same-verb pairs should be phonologically more similar than cognate-verb pairs). In fact, the correlation coefficient between verb status and phonological similarity was .83. Thus the greater structural priming effect observed with same verbs than with cognate verbs could also be attributed to the greater phonological similarity in same verbs than in cognate verbs. To determine which factor genuinely gave rise to the effect, we modeled the data with both factors as predictors and determined whether one predictor could be subsumed by the other, for example whether phonological similarity would be subsumed by verb status, such that verb status made an extra contribution to the model fit of the data independently of phonological similarity.

We therefore modeled the data from same-meaning verbs (same verbs and cognate verbs) across the two experiments, with verb status and phonological similarity as fixed predictors and priming (primed vs. unprimed response) as the dependent variable. We used the same method as we did in Experiments 1 and 2 to build LME models. We first added a predictor (e.g., phonological similarity) to see whether it significantly improved the fit of the model. Then we added the second predictor (e.g., verb status) to see whether it improved the model fit over and above the first predictor. If it did, we then concluded that first predictor could be subsumed by the second predictor. We also reversed the order of predictor addition (e.g., adding in verb status first and then phonological similarity) (see Table 5). As shown in Table 5, though phonological similarity and verb status alone significantly improved the model fit, phonological similarity did not improve the model fit over verb status, whereas verb status improved the model fit over phonological similarity. These results suggest that phonological similarity could be subsumed by verb status but verb status could not be subsumed by phonological similarity. Thus the apparent effect of phonological similarity could be simply attributed to verb status.

To further test whether phonological similarity had an effect in structural priming, we also modeled the priming data from cognate verbs only. If phonologically more similar cognate verbs (e.g., \textit{song–song}; "give") induce more structural priming than phonologically less similar cognate pairs (e.g., \textit{huan–wan}; "return"), then we would expect phonological similarity for a cognate pair to significantly predict the priming patterns in cognate verbs. We in fact did not obtain such a result: Phonological similarity did not significantly predict priming (Estimate = .05, SE = .06, \(z = .85, p > .1\)). In other words, phonologically more similar cognate verbs did not lead to more priming effects than less similar ones.\(^6\)

These findings indicate that differences in phonological similarity between cognate verbs and between same verbs cannot explain why the cognate boost to priming was smaller than the lexical boost to priming. Moreover, they indicate that phonological similarity of cognate verbs did not predict priming; in other words, it is not the case the phonologically more similar cognate pairs induced more structural priming than the phonologically less similar cognate pairs. An explanation that is more compatible with these findings is that cognate verbs have distinct lemmas (unlike same verbs), and that structural priming is based upon activation of these lemmas.

**General discussion**

Two experiments demonstrated both within- and between-language priming for two closely related languages, Mandarin and Cantonese. They also showed a general within-language over between-language advantage in structural priming when verbs were unrelated (though only when Mandarin was the response language). Critically, however, they additionally found a stronger boost for sentences involving the same verb (i.e., the lexical boost) than for sentences involving cognate verbs (i.e., the cognate boost). Further analyses showed that the difference in the lexical and cognate boosts to priming could not be attributed to a phonological similarity effect. Hence we conclude that it was a representational difference at the lemma level between same verbs and cognate verbs that led to the different boosts; specifically, repeated use of cognate verbs led to a weaker boost to priming than repeated use of same verbs because repeated use of cognate verbs involved activating distinct lemma representations, whereas repeated use of same verbs involved activating the same lemma representation. Our results thus support the separate lemma model, in which Cantonese–Mandarin bilinguals have different lemmas for Mandarin and Cantonese translation-equivalents. This suggests that even for languages as close as Mandarin and Cantonese, speakers differentiate cognates at the lexico-syntactic level.

The representation of lexico-syntactic information in bilinguals

Our experiments provide support for the separate lemma account, and are therefore broadly compatible with the proposal that bilinguals collectively represent common syntactic information between their two languages but maintain separate lexical (e.g., lemma) representations for translation-equivalents (Hartsuiker & Pickering, 2008; Schoonbaert et al., 2007). Importantly, this holds when

\(4\) Analyses for each experiment individually yielded similar patterns of results. In both analyses, phonological similarity was subsumed by verb status in explaining the priming differences between same verbs and cognate verbs; also, for cognate verbs, phonological similarity did not significantly predict priming effects.
the translation-equivalents are cognates, as in our experiments.

The conclusion that cognates have separate lemmas appears to contradict recent ERP evidence showing a cognate facilitation effect early in lexical processing, which could be interpreted as support for a shared lemma representation for cognates (Strijkers et al., 2010). However, Strijkers et al. speculated that the cognate facilitation effect could actually be a frequency effect in disguise: Because cognates are similar in form, the processing of any cognate word would then activate its counterpart via feedback, increasing cognate words’ overall frequency (see also Lalor & Kirsner, 2001). Such an account is consistent with our finding that cognates do not have a shared lemma. Such feedback might in principle arise at either at a morphological level (e.g., Davis et al., 2010; Lalor & Kirsner, 2001; Sanchez-Casas & Garcia-Albea, 2005) or a phonological level (Costa et al., 2000; Costa et al., 2005). Our finding that the cognate boost is not affected by phonological similarity suggests that the morphological level might be a more likely locus for this effect.

Our data also relate to the question of within-versus-between-language priming for unrelated verbs. Previous studies have found similar priming in these cases (Kantola & van Gompel, 2011; Schoonbaert et al., 2007). In contrast, Experiment 1 showed stronger within- than between-language priming with unrelated verbs. However, the trend toward stronger within- than between-language priming in Experiment 2 was not significant.

Hartsuiker and Pickering (2008) assumed a language control mechanism which is executed via the language nodes, but they did not assume that the language nodes act like other nodes (see also Hartsuiker et al., 2004; Schoonbaert et al., 2007). Specifically, they assumed that activation from a lemma to a conceptual node leads to activation of other lemmas linked to that conceptual node, but did not assume that from a lemma to a language node leads to activation of other lemmas linked to that language node. They therefore predicted similar within- and between-language priming for unrelated verbs.

But the results of Experiment 1 support the alternative account in which language nodes act like other nodes. If so, activation of a language node leads to activation of all lemmas of that language, thus leading to a within-language boost for unrelated verbs. This mechanism is consistent with the proposal that lemmas in a bilingual lexicon are tagged for language (Green, 1998), with the speaker selecting a language and in doing so activating the relevant “language task schema” (similar to the language node). This language selection then inhibits the activation of lexical information of the non-response language and ensures the use of lexical items of the desired language (see Dijkstra & Van Heuven, 1998; Dijkstra & Van Heuven, 2002, for a similar inhibition mechanism for bilingual visual word recognition). Language-selection presumably takes place “early,” when the speaker is formulating the message. This leads to a much greater activation of lemmas that are linked to that language node (i.e., lemmas of that language) than lemmas that are not linked to that language node. For instance, the processing of a prime containing di (“pass”) in a DO form would not only activate the lemmas (e.g., DI) and combinatorial node (DO) in the prime, but also the language node MANDARIN. The activation of the MANDARIN node would then activate all Mandarin lemmas, including huan (Mandarin “return”), but not wan (Cantonese “return”). Coactivation of huan and the DO node would strengthen the activation of the link between them (in comparison to the link between wan and the DO node). Thus, speakers would be more likely to use DO when the target is huan than when the target is wan, as found in Experiment 1. Note that a similar mechanism has been used to explain the semantic boost in monolingual production (Cleland & Pickering, 2003). However, we reiterate that Experiment 1 provides support for this account, but Experiment 2 does not.

This account is compatible with Green’s model, except that his model is couched in terms of inhibition of the non-selected language rather than facilitation of the lemmas of the selected language. It may explain why bilinguals tend to have very strong control over which language(s) to use (e.g., Grosjean, 1997) – they rarely produce a word from the wrong language – but can explain occasional misselection of lemmas from the inappropriate language, in that an unintended language node will occasionally accrue sufficient activation from other sources to override activation from the language node.

The nature of between-language structural priming

Schoonbaert et al. (2007) argued the translation-equivalence boost occurs during the processing of the target response rather than during the processing of the prime response (e.g., Cleland & Pickering, 2003). They proposed this account to explain their finding of a translation-equivalence boost from their L1 (Dutch) to their L2 (English) but not vice versa. Because models of bilingual lexical processing (e.g., Kroll & Stewart, 1994) suggest weaker links from concepts to L2 lemmas than to L1 lemmas, the lack of a translation-equivalence boost from L2 to L1 could be due to the target L1 lemma activating the shared concept, but the shared concept failing to activate the L2 lemma (see Schoonbaert et al., 2007, pp. 165–166). However, our study showed a (cognate) translation-equivalence boost both from Cantonese to Mandarin, and from Mandarin to Cantonese. (Recall that our participants were native Cantonese speakers, though they had used Mandarin from early childhood.) It is therefore possible that Schoonbaert et al. simply failed to detect a real effect, and so it may be premature to assume a target-based account of the translation-equivalence boost, or indeed of the lexical boost within a language.

Note that the numerical magnitude of priming was quite different in our study from Schoonbaert et al. (2007). As they measured priming effects in terms of the proportion of PO responses out of both DO and PO responses and interpreted their data primarily in terms of the lexical boost and the translation-equivalence boost, we do the same here. Our Experiment 1 showed a .38 lexical boost (.96–.30) – (.82–.54); i.e., the difference in terms of PO proportion between the PO and DO primes in the within-language same-meaning verb condition minus that in the within-language different-meaning verb condition] and a .28 translation-
equivalence boost (see Table 1), and Experiment 2 showed a .28 lexical boost and a .12 translation-equivalence boost (see Table 3). In contrast, Schoonbaert et al. showed lexical boosts of .27 and .29 in Experiments 1 and 3, and translation-equivalence boosts of .09 and .01 (which was non-significant) in Experiments 2 and 4. When considered as a whole (i.e., collapsed over experiments), the lexical boost in our experiments was 1.65 times the (cognate) translation-equivalence boost; but the lexical boost in Schoonbaert et al. was 7.00 times the translation-equivalence boost.

One possible explanation for the difference between the two studies is that cognate verbs (as used in our study) might give rise to a greater boost than noncognate verbs (as mainly used in Schoonbaert et al.’s [2007] study) do. In other words, the cognate translation-equivalence boost would be larger than the noncognate translation-equivalence boost. Of course, it is still an open question whether cognates and noncognates lead to different boosts and it is not an attempt for our study to address this question. There are also other possible accounts for the disparity between our study and Schoonbaert et al.’s study. For instance, apparent translation-equivalents in English and Dutch might have some differences in meaning, with these differences reducing the magnitude of priming. Alternatively, the large translation-equivalence boost may have occurred because our participants were fairly balanced bilinguals who acquired both languages early, or because they tended to switch languages more regularly than Schoonbaert et al.’s participants.

The relationship between Mandarin and Cantonese

Linguists (such as sociolinguists) have long been interested in whether language variants are languages or dialects of a single language (e.g., Haugen, 1966). In contrast, psycholinguists have shown little interest in the representation and processing of languages vs. dialects, and work on bilingualism has almost entirely focused on speakers of clearly distinct languages. However, our results provide some evidence that bears on this issue. The stronger lexical boost than cognate boost to priming suggests that Mandarin and Cantonese have separate lemmas associated with different language nodes. Furthermore, the combined analyses showed that the PO construction is more frequently used in Cantonese than in Mandarin and the DO construction is more frequently used in Mandarin than in Cantonese (when participants described the same event). These differences suggest that speakers of Mandarin and Cantonese not only represent cognate verbs differently, but also differ in the relative strengths of the links that connect verbs to the DO and PO nodes, with a stronger link to the DO node in Mandarin than in Cantonese, and a stronger link to the PO node in Cantonese than in Mandarin. We therefore have strong evidence that Mandarin and Cantonese have distinct lemma representations. Their phonological differences, and the fact that they are mutually unintelligible (Tang & van Heuven, 2009), argue further for at least partially distinct phonological representations. In all, then, we have evidence that Mandarin and Cantonese are represented separately, just as (uncontroversially distinct) languages such as English and Dutch are. Hence our experiments suggest that they have at least some psycholinguistic characteristics associated with different languages rather than different dialects.

In conclusion, two experiments demonstrated that cognate verbs induced a smaller boost to priming than same verbs did, irrespective of phonological similarity. These results suggest that even in very closely related languages, cognate verbs have separate lemma representations rather than a shared lemma representation. Overall, our results support a model of bilingual lexico-syntactic representation in which speakers of closely related languages represent lexico-syntactic information in a similar way to speakers of clearly distinct languages.

Acknowledgements

We would like to thank Martin Corley, Patrick Sturt, Janet McLean, Mateo Obregón, Michael Jin and Yanping Dong for help with statistics and/or experiments. The research was supported by an Edinburgh University studentship and an ORS award (to Z. G. C.) and ESRC Grant No. RES-062-23-0376 (to M. J. P.). We also acknowledge the support by the Center of Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, China during the experiments.

A. Experimental materials for Experiments 1 and 2

In Set 1, the first sentence corresponds to the DO (same-meaning/different-meaning verb) prime and the second sentence to the PO (same-meaning/different-meaning verb) prime. The last sentence is a DO description of the target picture. From Set 2 onward, we only provide the DO (same-meaning/different-meaning verb) prime sentence and the target picture description in the DO form. The same materials were used in Experiments 1 and 2. All the prime sentences had a Mandarin version and a Cantonese version.

<table>
<thead>
<tr>
<th>No.</th>
<th>English / Mandarin / Cantonese sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>牛仔送[给]给水手一本书。 (The cowboy gave/donated the sailor a book.)</td>
</tr>
<tr>
<td></td>
<td>(DO, same-verb/different-verb)</td>
</tr>
<tr>
<td></td>
<td>牛仔送[给]给了一本书给水手。 (The cowboy gave/donated a book to the sailor.)</td>
</tr>
<tr>
<td></td>
<td>(PO, same-verb/different-verb)</td>
</tr>
<tr>
<td></td>
<td>警察送给士兵一顶帽子。 (The policeman gave the soldier a hat.)</td>
</tr>
<tr>
<td></td>
<td>(Target)</td>
</tr>
<tr>
<td>2.</td>
<td>公主还[给]给牧师一个排球。 (The princess returned/gave the priest a volleyball.)</td>
</tr>
<tr>
<td></td>
<td>教授还给游泳选手一条香蕉。 (The professor returned the swimmer a banana.)</td>
</tr>
<tr>
<td>3.</td>
<td>海盗让[给]给水手一个蛋糕。 (The pirate let-have returned the sailor a cake.)</td>
</tr>
<tr>
<td></td>
<td>厨师让给拳击手一个排球。 (The chef let-have the boxer a volleyball.)</td>
</tr>
<tr>
<td>4.</td>
<td>海盗卖[给]给水手一个杯子 (The pirate sold/let-have the sailor a cup.)</td>
</tr>
<tr>
<td></td>
<td>警察卖给拳击手一把枪。 (The policeman sold the boxer a gun.)</td>
</tr>
</tbody>
</table>
1. 修女/卖给医生一个水壶。(The nun awarded/sold the doctor a jug.) 画家/给小丑一个苹果。(The artist awarded the clown an apple.)
2. 教授/卖给消防员一条香蕉。(The professor bought/ awarded the fireman a banana.) 修女/卖给士兵一本书。(The nun bought the soldier a book.)
3. 服务员/卖给水手一把枪。(The waitress lent/bought the sailor a gun.)
4. 警察/给小偷一条香蕉。(The policeman returned the thief a banana.)
5. 画家/卖给医生一个水壶。(The artist awarded the clown a jug.)
6. 教授/卖给消防员一条香蕉。(The professor bought/ awarded the fireman a banana.) 修女/卖给士兵一本书。(The nun bought the soldier a book.)
7. 服务员/卖给水手一把枪。(The waitress lent/bought the sailor a gun.)
8. 印第安人/卖给医生一个水壶。(The Indian passed/lent the doctor a jug.)
9. 牛仔/卖给水手一条香蕉。(The cowboy rented/passed the sailor a banana.)
10. 警察/给游泳选手一位帽子。(The policeman rented the swimmer a hat.)
11. 警察/给拳击手一个蛋糕。(The policeman handed the boxer a cake.)
12. 画家/卖给医生一本书。(The artist passed the doctor a book.)
13. 牛仔/卖给小丑一条香蕉。(The cowboy rented/passed the clown a banana.)
14. 厨师/卖给水手一个蛋糕。(The chef gave/submitted the sailor a cake.)
15. 家庭/卖给士兵一个苹果。(The artist returned/submitted the soldier a banana.)
16. 画家/给水手一个苹果。(The artist let-have/gave the sailor an apple.)
17. 警察/卖给士兵一个水壶。(The policewoman sold/ returned the soldier a jug.)
18. 厨师/卖给拳击手一顶帽子。(The chef awarded/let-have the boxer a hat.)
19. 厨师/卖给游泳选手一个水壶。(The chef bought/sold the swimmer a jug.)
20. 海盗/卖给小偷一把枪。(The pirate lent/awarded the thief a gun.)
21. 厨师/卖给舞蹈演员一个西瓜。(The chef passed/bought the dancer a watermelon.)
22. 画家/卖给医生一顶帽子。(The artist lent/lent the doctor a hat.)
23. 厨师/卖给修女一个杯子。(The chef threw/passed the nun a cup.)
24. 牛仔/卖给游泳选手一个苹果。(The cowboy handed/ rented the swimmer an apple.)
25. 画家/卖给游泳选手一个蛋糕。(The artist passed/ gave the swimmer a cake.)
26. 厨师/卖给士兵一把枪。(The professor sold/gave the soldier a gun.)
27. 修女/卖给医生一本书。(The nun bestowed-upon/ sold the doctor a book.)
28. 修女/卖给小丑一条香蕉。(The nun donated/bought the clown a banana.)
29. 画家/卖给游泳选手一个蛋糕。(The artist passed the dancer a cake.)
30. 厨师/卖给修女一顶帽子。(The professor donated/handed the priest a hat.)
31. 画家/卖给警察一本书。(The artist donated the policeman a book.)
32. 公主/卖给士兵一本书。(The princess submitted/donated the soldier a book.)
33. 服务员/卖给舞蹈演员一个蛋糕。(The waitress submitted the dancer a cake.)
B. Ratings of phonological similarity (1 = very alike, 7 = very different)

<table>
<thead>
<tr>
<th>Verbs</th>
<th>Average rating (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>递 (give)</td>
<td>1.88 (1.22)</td>
</tr>
<tr>
<td>还 (return)</td>
<td>4.1 (1.86)</td>
</tr>
<tr>
<td>让 (let-have)</td>
<td>4.7 (1.85)</td>
</tr>
<tr>
<td>卖 (sell)</td>
<td>2.2 (1.31)</td>
</tr>
<tr>
<td>奖 (award)</td>
<td>3.88 (1.65)</td>
</tr>
<tr>
<td>买 (buy)</td>
<td>2.07 (1.42)</td>
</tr>
<tr>
<td>借 (lend)</td>
<td>2.36 (1.2)</td>
</tr>
<tr>
<td>递 (pass)</td>
<td>3.94 (1.46)</td>
</tr>
<tr>
<td>租 (rent)</td>
<td>4 (1.75)</td>
</tr>
<tr>
<td>抛 (throw)</td>
<td>1.85 (1.86)</td>
</tr>
<tr>
<td>拿 (hand)</td>
<td>2.97 (2.26)</td>
</tr>
<tr>
<td>交 (submit)</td>
<td>4 (1.78)</td>
</tr>
<tr>
<td>赠 (bestow-upon)</td>
<td>2.88 (1.81)</td>
</tr>
<tr>
<td>押 (donate)</td>
<td>4.85 (1.57)</td>
</tr>
</tbody>
</table>

References


