



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Does language similarity affect representational integration?

**Citation for published version:**

Huang, J, Pickering, M, Chen, X, Cai, Z, Wang, S & Branigan, H 2019, 'Does language similarity affect representational integration?', *Cognition*, vol. 185, pp. 83-90. <https://doi.org/10.1016/j.cognition.2019.01.005>

**Digital Object Identifier (DOI):**

[10.1016/j.cognition.2019.01.005](https://doi.org/10.1016/j.cognition.2019.01.005)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Cognition

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



Does language similarity affect representational integration?

Jian Huang<sup>1</sup>, Martin J. Pickering<sup>2\*</sup>, Xuemei Chen<sup>1</sup>, Zhenguang Cai<sup>3</sup>, Suiping Wang<sup>1\*</sup>, Holly  
P. Branigan<sup>2</sup>

<sup>1</sup>School of Psychology, South China Normal University

<sup>2</sup>Department of Psychology, University of Edinburgh

<sup>3</sup>Department of Psychology, University of East Anglia

\*Corresponding author

Martin Pickering

Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, United Kingdom

Email: [martin.pickering@ed.ac.uk](mailto:martin.pickering@ed.ac.uk)

Suiping Wang

School of Psychology, South China Normal University, Guangzhou, 510631, China

Email: [wangsuiiping@m.scnu.edu.cn](mailto:wangsuiiping@m.scnu.edu.cn)

RUNNING HEAD: Language similarity

## **Abstract**

Previous studies have suggested that multilingual speakers do not represent their languages entirely separately but instead share some representations across languages. To determine whether sharing is affected by language similarity, we investigated whether participants' tendency to repeat syntax across languages was affected by language similarity. In three cross-linguistic structural priming experiments, trilingual Mandarin-Cantonese-English participants heard a sentence in Cantonese or English (which they matched to a picture) and then described a dative event in Mandarin. When prime and target sentences involved different actions (Experiment 1), structural priming was unaffected by language similarity. But when prime and target involved the same action (Experiments 2 and 3), priming was stronger between related languages (i.e., Cantonese to Mandarin) than unrelated languages (i.e., English to Mandarin). Similar languages are not more integrated than dissimilar languages overall, but the representations that connect lexical and syntactic information are more closely integrated.

*Keywords:* Multilingualism, Syntax, Structural Priming, Language, Cantonese, Mandarin

More than half the world's population speaks more than one language (Grosjean, 1982). But some combinations of languages are popularly believed to be harder to learn than others. People are often impressed to discover someone speaks proficient Mandarin and English, less so when they speak proficient Mandarin and Cantonese. One possible reason is the overall similarity between these language pairs: Although every language has its own distinct sound patterns, vocabulary, and grammar, languages may share some characteristics to a greater or lesser degree. Mandarin and Cantonese share many properties (like many other Chinese languages such as Hakka and Chaoshanese), including a large number of cognates (words with the same meaning and similar pronunciation, e.g., *di/dai* 'to pass'), many grammatical structures (e.g., the *Double Object* (DO) structure: *Niuzai di-gei shuishou yitiao xiangjiao*/*Ngaozai dai-bei suisau yattiu heungjiu*, 'the cowboy passed the sailor a banana'; the *Ba* structure *Niuzai ba shuishou da-le*/*Ngaozai ba suisau daa-zo*, 'the cowboy hit the sailor'), considerable phonology (e.g., a tonal system, syllable structure), and orthography (i.e., characters), whereas Mandarin and English share substantially fewer properties, for example having few (if any) cognates, sharing the DO (*Niuzai di-gei shuishou yitiao xiangjiao*/*The cowboy passed the sailor a banana*) but not the *Ba* structure, and using different orthography. The US Foreign Service Institute suggests that the substantial linguistic (and cultural) differences between Mandarin and English make the former 'exceptionally difficult' for native English speakers to become proficient in, compared to 'closely related' languages such as Dutch (2200 vs. 675 hours; "Language Assignments", 2015). More generally, it

assumes that similarity to their native language (L1) affects how easily multilingual (including bilingual) speakers acquire a new language (L2).

Does language similarity also affect how multilingual speakers represent their languages? Many studies have suggested that multilinguals do not represent their languages entirely separately, but instead share, or integrate, some representations across languages (e.g., regarding cognates) (De Bot, 1992; Hartsuiker, Pickering, & Veltkamp, 2004). In this paper, we focus on the representation of syntax and associated lexical information, and ask whether multilingual Mandarin-Cantonese-English speakers (who also speak Hakka or Chaoshanese) share more information between Mandarin and Cantonese than between Mandarin and English. Language similarity can be defined typologically (e.g., with respect to word orders or word complexity) or historically/genetically (e.g., Indo-European vs. Sino-Tibetan), so we chose languages for which all measures of similarity are in agreement: Mandarin and Cantonese (which are unambiguously close) versus Mandarin and English (which are unambiguously distant).

To investigate whether language similarity affects syntactic integration, we consider cross-linguistic structural priming from Cantonese and English to Mandarin. Following Bock (1986), we know that people tend to repeat aspects of the structure of utterances that they have recently produced or comprehended, including syntactic structure (e.g., using a Prepositional Object [PO] structure such as *The girl gives the paintbrush to the man* more frequently after another PO than after a DO). These

findings are used to understand both the mechanisms involved in language processing and the nature of linguistic representations themselves (Branigan & Pickering, 2017). Structural priming occurs across types of structures (e.g., actives vs. passives, types of noun phrases), in native and non-native speakers, and in all languages that have been tested (Pickering & Ferreira, 2008). But most strikingly, it occurs between languages in multilinguals, with speakers who encounter a structure in one language showing an increased likelihood of using an equivalent structure in another language (i.e., a structure involving the same phrasal categories in the same order; e.g., Hartsuiker et al., 2004; Loebell & Bock, 2003).

These results suggest that bilinguals have language-general integrated syntactic representations for some structures. For example, hearing a passive in Spanish activates a passive representation that is shared between Spanish and English, making it more likely that the speaker will subsequently produce a passive when using English. Cross-linguistic priming (e.g., from Swedish to English) is often as strong as within-language priming (e.g., from English to English; see Van Gompel & Arai, 2017), suggesting that multilinguals can have fully integrated representations for equivalent structures in their different languages. But does the degree of integration differ depending on language similarity? For example, do speakers who are proficient in Mandarin, Cantonese, and English (henceforth, *Mandarin-Cantonese-English speakers*; note that they also speak Hakka or Chaoshanese) integrate their representations for equivalent structures in Mandarin and English to the same extent

as for equivalent structures in Mandarin and Cantonese?

It is possible that multilingual speakers develop more highly integrated representations for languages that are more similar overall than for languages that are less similar. If so, multilingual Mandarin-Cantonese-English speakers who are similarly proficient in Cantonese and English would have more integrated representations for equivalent structures such as the DO structure in Mandarin and Cantonese than in Mandarin and English. Thus they might activate a shared representation of the DO whenever they process a DO sentence in Cantonese or English, and this activation would affect their subsequent choice of whether to use a DO in Mandarin, leading to cross-linguistic structural priming effects – but crucially, these effects would be consistently stronger for Cantonese than for English. This account is compatible with evidence that between-language priming can be smaller than within-language priming (Cai et al., 2011).

Alternatively, multilingual speakers might develop equally integrated representations for languages, irrespective of the languages' overall similarity. If so, multilingual Mandarin-Cantonese-English speakers would develop equally integrated representations for equivalent structures such as the DO or PO in Mandarin, Cantonese, and English. Thus they might activate a shared representation of the DO whenever they process a Mandarin DO sentence, and this activation would lead to cross-linguistic structural priming effects that would be equally strong in Cantonese and English. This account is compatible with other evidence that between- and

within-language priming can be equivalent (e.g., Kantola & Van Gompel, 2011; Schoonbaert et al., 2007).

We report three structural priming experiments that investigated syntactic integration in multilingual Mandarin-Cantonese-English speakers. In Experiment 1, participants listened to Cantonese and English PO and DO sentences, and then produced Mandarin descriptions of ditransitive events involving a different action. We investigated whether they were more likely to repeat syntax after a Cantonese prime than an English prime, suggesting that language similarity affects syntactic representation, or equally likely to repeat syntax after a Cantonese prime as an English prime, suggesting that language similarity does not affect syntactic representation. Experiments 2 and 3 tested whether similarity affects the integration of syntactic and lexical representations, by having participants comprehend Cantonese and English PO and DO sentences, and then produce Mandarin descriptions of ditransitive events involving the same action.

## **Experiment 1**

### **Method**

*Participants.* 32 participants (8 male; aged 19-24 years with an average of 21) were paid 25 RMB to take part. We used participants who spoke Mandarin as their dominant language (acquired during kindergarten alongside their mother tongue Hakka or Chaoshanese) and subsequently learned Cantonese and English. We confirmed participants' language background and proficiency through a questionnaire

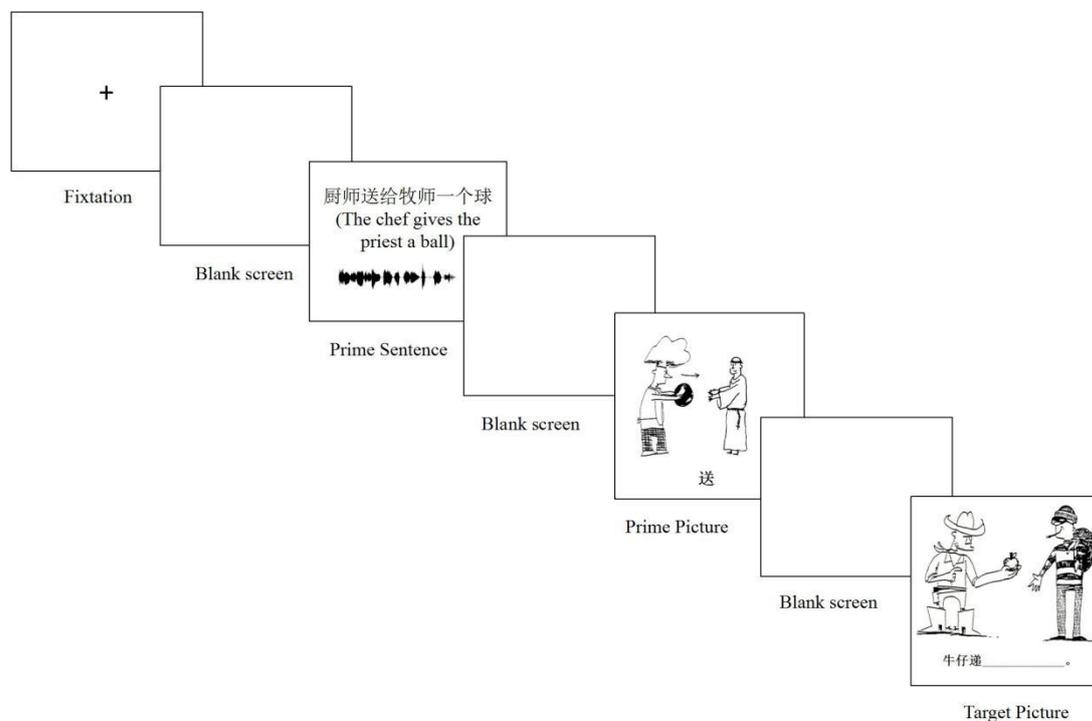
in which they reported age of acquisition for Mandarin/Cantonese/English, and rated their general proficiency and sub-categories of proficiency on a 5-point scale (1 = very poor, 5 = very proficient) in Cantonese and English (see Table 1). Participants reported acquiring Mandarin before Cantonese, and Cantonese before English, and speaking Mandarin with greater overall proficiency than either Cantonese or English (*t*-tests, all *ps* < .05). Their self-ratings for Cantonese and English proficiency did not differ (all *ps* > .1), except for age of acquisition (*t* = -3.13, *p* = .004) and listening comprehension (*t* = 3.05, *p* = .005).

*Materials.* The 32 sets of experimental and 96 filler items were based on Cai et al. (2011) (for materials see <https://osf.io/znk37/>). Each item consisted of a (spoken) prime sentence, a prime picture, a target picture, and target preamble (see Fig. 1). For the experimental items, prime sentences had four versions yielded by crossing Prime Structure (DO vs. PO) and Prime Language (Cantonese vs. English) (e.g., *Cyusi dai-bei muksi jatgo kau*, ‘chef gives-to LE priest a ball’; *Cyusi dai-zo kau bei muksi*, ‘chef give-LE ball to priest’; *The chef gives the priest a ball*; *The chef gives a ball to the priest*).

Table 1. Language background self-ratings in Experiments 1-3 (Standard Deviation in brackets)

	Experiment 1			Experiment 2			Experiment 3		
Language background	Mandarin	Cantonese	English	Mandarin	Cantonese	English	Mandarin	Cantonese	English
Listening Comprehension	9.00 (1.18)	6.75 (1.55)	5.81 (1.20)	9.34 (2.19)	6.41 (2.00)	5.94 (1.16)	8.96(0.95)	6.82(1.80)	6.19(1.58)
Reading Comprehension	8.59 (2.41)	6.25 (1.98)	6.16 (1.48)	9.47 (0.76)	6.88 (1.91)	6.72 (1.30)	9.22(0.86)	6.85(1.91)	6.99(1.65)
Speaking Fluency	8.91 (1.09)	5.78 (2.31)	5.69 (1.67)	8.94 (0.98)	5.41 (2.20)	5.84 (1.59)	9.00(1.04)	6.12(2.03)	6.04(1.79)
Speaking Pronunciation	7.84 (1.46)	5.25 (2.05)	5.53 (1.72)	8.44 (1.13)	5.22 (2.24)	5.88 (1.54)	8.26(1.23)	5.79(2.06)	6.16(1.87)
General Proficiency	4.44 (0.72)	3.03 (0.86)	2.97 (0.82)	4.75 (0.44)	2.84 (1.14)	3.09 (0.64)	4.57(0.58)	3.16(1.02)	3.29(0.93)
Age of acquisition	5.94 (1.61)	7.97 (3.46)	9.69 (2.15)	4.97 (2.88)	8.88 (5.42)	9.00 (2.18)	4.35(2.60)	6.78(3.17)	8.82(2.57)

Fig. 1. Example experimental trial.



Note. The character below the prime picture means “give”; and the characters below the target picture mean “The cowboy pass\_\_\_\_\_.” In Experiments 1 and 2, the target preamble was presented visually as depicted; in Experiment 3, it was presented auditorily.

The prime picture depicted a ditransitive event, with the positions of the agent and the recipient (e.g., agent on the left and recipient on the right) counter-balanced across items; the theme always appeared centrally. The prime picture had a verb corresponding to the depicted action (a character for Cantonese primes and an English word for English primes) printed underneath. Half of the prime pictures matched the

prime sentence in meaning; half had a mismatching agent, recipient, or theme.

The target picture depicted a ditransitive event unrelated to the prime sentence/picture, and with no overlapping event participants (agent, recipient, or theme). The positions of the agent and recipient were counter-balanced across items. A sentence preamble (a noun phrase followed by a verb) appeared underneath the picture in Chinese characters; this could be continued as a DO or PO (but not a *ba-* or *bei-*construction) (e.g., *Niuzai di*, cowboy pass; ‘the cowboy passed...’).

In the filler items, primes comprised 24 DO sentences (12 Cantonese and 12 English, intended to boost Mandarin DO responses; see Cai et al., 2011), together with 21 intransitive and 51 transitive sentences. Half the prime pictures matched the prime sentence, and the other half mismatched in one event participant. Target pictures comprised 39 intransitive pictures (with a preamble comprising the agent) and 57 transitive pictures (with a preamble comprising the agent and verb).

A female speaker from the same population as the participants digitally recorded the prime sentences. We created 4 lists using a Latin-square design, each containing 8 Cantonese DO, 8 Cantonese PO, 8 English DO, and 8 English PO primes, in addition to the 96 filler items. Items were presented in two blocks (Cantonese vs. English primes), with block order counter-balanced across participants, and item order within blocks individually randomized.

*Procedure.* Participants were tested individually using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Following Cai et al. (2011), we first familiarized

participants with the names of the experimental entities. They were then randomly assigned to one of the 4 lists. Each block began with 3 practice trials. Each trial began with a fixation cross (500 ms) followed by a blank screen (200 ms), then the prime sentence played via headphones, immediately followed by the prime picture. Participants were instructed to press the “F” key if the prime sentence and picture matched, and the “J” key if they did not, triggering a blank screen (200 ms), followed by the target picture. Participants described the target picture in Mandarin using a full sentence that began with the given sentence preamble; responses were digitally recorded. Participants pressed the spacebar to trigger the next trial. The experiment lasted approximately 45 minutes.

*Scoring.* Following Cai et al. (2011), target descriptions were scored as DO if the verb in the preamble was followed by a noun phrase corresponding to the recipient and then by a noun phrase corresponding to the theme (e.g., *Niuzai di-gei xiaotou yige pingguo*, ‘The cowboy passes the thief an apple’); as PO if the verb was followed by a noun phrase corresponding to the theme and then a prepositional phrase corresponding to the preposition *gei* encoding the recipient (e.g., *Niuzai di-le yige pingguo gei xiaotou*, ‘The cowboy passes an apple to the thief’); all other responses were coded as Other, and excluded from analysis. For the sake of statistical analyses, we further coded a response as primed if it had the same structure as the prime (e.g., DO response after a DO prime) or unprimed if it did not have the same structure as the prime (e.g., a DO response following a PO prime); see below.

## Results

The data and statistical analyses for this and the following experiments are available at <https://osf.io/znk37/>. We used logistic mixed effects (LME) modelling to analyze PO and DO responses (see Table 1 for frequency by condition). Following Cai et al. (2011), our analyses compared the likelihood of a response being a primed or unprimed response following a particular prime structure. This way, a structural priming effect would manifest as a significant intercept (i.e., whether there were more primed than unprimed responses), and a modulation of this effect depending on prime language would manifest as a significant effect of prime language (i.e. whether there was more priming for one prime language than the other). The LME model included prime language (z-score transformed) as a fixed effect. We adopted the maximal random effect structure justified by the design (Barr, Levy, Scheepers, & Tily, 2013).

As we used LME analyses in our experiments, we used the R package SIMR (Green & MacLeod, 2016), which allows for power analysis for LME models, to determine whether our experiments were sufficiently powered to detect priming from Cantonese/English to Mandarin, under different-meaning (Experiment 1) or same-meaning verb condition (Experiments 2 and 3). To estimate the most likely effect size of priming from the literature, we resorted to Cai et al. (2011), which the current experiments modelled after in terms of design, syntactic constructions, materials, and statistical analyses. From there we calculated the priming effect from Cantonese to Mandarin to be 0.33 in the different-meaning verb condition and 1.07 in

the same-meaning verb condition (see Cai et al., 2011, p.438, Table 2). Using these figures for the SIMR power analysis revealed that all the three experiments reported in the paper have almost 100% statistical power to detect cross-language priming either under the different-meaning verb or same-meaning verb condition (indeed when we used only half of these effect sizes in the simulation, our experiments also have more than 80% statistical power to detect cross-language priming in either meaning conditions).

Table 2 presents the proportion of priming as a function of prime language. There was a significant intercept ( $\beta = 0.18$ ,  $SE = 0.06$ ,  $z = 2.89$ ,  $p = .004$ ), supporting a structural priming effect: In their target descriptions, participants were more likely to use the syntactic structure used in the prime sentence than to use the alternative structure. Cross-language priming occurred both from Cantonese to Mandarin ( $\beta = 0.18$ ,  $SE = 0.09$ ,  $z = 1.96$ ,  $p = .050$ ), and from English to Mandarin ( $\beta = 0.19$ ,  $SE = 0.09$ ,  $z = 2.13$ ,  $p = .034$ ).

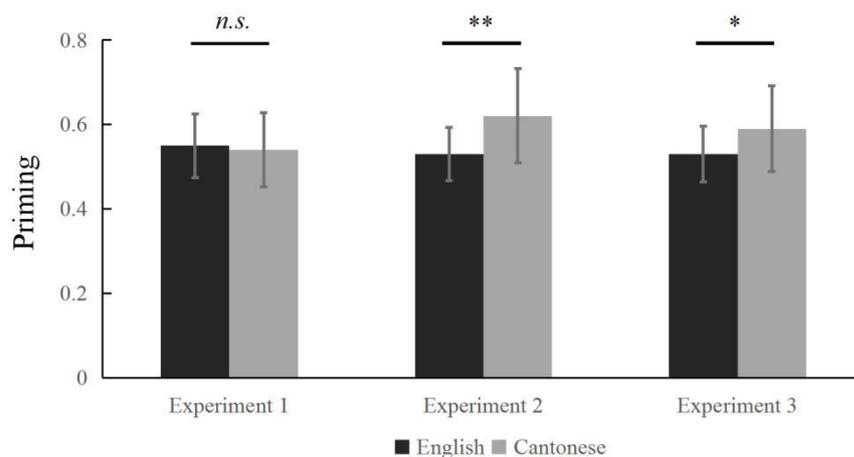
Table 2. Frequency of target responses and priming effect by prime condition in Experiments 1-3, where the priming effect is calculated as the difference in the proportion of primed and unprimed responses.

	Experiment 1				Experiment 2				Experiment 3			
	Eng. DO	Eng. PO	Can. DO	Can. PO	Eng. DO	Eng. PO	Can. DO	Can. PO	Eng. DO	Eng. PO	Can. DO	Can. PO
DO	66	41	65	44	96	78	118	55	152	116	186	98
PO	189	212	186	209	159	175	137	198	355	388	322	409
Other	1	3	5	3	1	3	1	3	5	8	4	5
DO prop	0.26	0.16	0.26	0.17	0.38	0.31	0.46	0.22	0.30	0.23	0.37	0.19
Primed	278		274		271		316		540		595	
Unprimed	230		230		237		192		471		420	

Priming	0.55	0.54	0.53	0.62	0.53	0.59
---------	------	------	------	------	------	------

Critically, prime language was not a significant predictor ( $\beta = -0.007$ ,  $SE = 0.06$ ,  $z = -0.12$ ,  $p = .909$ ): Participants were primed to produce DO and PO structures in Mandarin to the same extent irrespective of whether the prime was in Cantonese or English (see Fig. 2). We subsequently tested how likely it was that the data reflected a null difference between prime languages by using Bayes factor (BF), which helps to quantify the relative likelihood of the null and alternative hypotheses concerning an effect on the basis of the observed data (Wagenmakers, 2007). Following Wagenmakers, we used the Bayesian information criterion (BIC) from an alternative-hypothesis LME model (i.e., with the fixed effect of prime language) and the BIC from a null-hypothesis LME model (without the fixed effect of prime language), and found that the null hypothesis (i.e., that the effect was genuinely absent) was 32 times ( $BF_{01} = 31.5$ ) more likely than the alternative hypothesis (i.e., that the effect was real but not detected).

Fig. 2. Priming effects in Experiments 1-3.



Note: \*  $p < .05$ , \*\*  $p < .01$ . Error bars reflect standard errors calculated for a by-participants

analysis.

Further LME analyses showed that the priming effects from Cantonese and from English were not significantly correlated with either age of acquisition or listening comprehension proficiency (the only two measures with significant between-language differences; all  $ps > .1$ ); in addition, the priming effects did not vary as a function of block order ( $ps > .1$ ).

## **Discussion**

Experiment 1 thus found equivalent cross-linguistic structural priming between similar (Cantonese-Mandarin) and dissimilar (English-Mandarin) languages. These results suggest multilinguals share purely syntactic representations across their languages, irrespective of language similarity.

However, previous research has shown that priming is stronger when a verb or its translation-equivalent are repeatedly used, suggesting that syntactic representations are closely linked to lexical representations (Pickering & Branigan, 1998; Cai et al., 2011). It is possible that language similarity does not affect the representation of purely syntactic information, but does affect the extent to which lexical and syntactic representations are integrated. To test this possibility, Experiment 2 investigated Cantonese-Mandarin and English-Mandarin cross-linguistic priming when prime and target sentences involved the same action, and hence used translation-equivalent verbs.

## Experiment 2

Experiment 2 used the same design as Experiment 1, but primes and targets were re-paired so that the prime and target sentences used translation-equivalent verbs.

### Method

*Participants.* 32 further participants (8 male; 18-23 years with an average of 20) from the same population as Experiment 1 were paid 25 RMB. As in Experiment 1, participants reported acquiring Mandarin at an earlier age than both Cantonese and English, and speaking Mandarin with greater overall proficiency than Cantonese and English (*t*-tests, all *ps* < .05). Their ratings for Cantonese and English did not differ (all *ps* > .1).

*Materials.* We took the materials from Experiment 1 and re-paired primes and targets to create new 32 sets of materials in which the Cantonese or English prime and the Mandarin target used verbs of the same meaning (i.e. cognate translation equivalents between Cantonese and Mandarin and non-cognate translation equivalents between English and Mandarin). In the six cases where the re-pairing resulted in one entity (e.g., the agent) being repeated between the prime and target, we replaced the repeated entity with a different entity. In the filler items, primes comprised 8 intransitive, and 22 transitive sentences with repeated verbs across prime and target, 24 DO, 13 intransitive, and 29 transitive sentences with different verbs across prime and target.

*Translation equivalence pre-test.* It is possible that cross-language priming could be influenced by the extent to which participants treated the relevant verbs as translation equivalents in Cantonese and Mandarin versus English and Mandarin. That is, participants might more consistently associate Cantonese prime verbs and Mandarin target verbs with the same concept (e.g., mapping *dai* and *di* to the same concept) than English prime verbs and Mandarin target verbs (e.g., mapping *pass* and *di* to the same concept). To test this possibility, we recruited 32 further participants from the same population as the main experiment. Half of the participants were presented with the 32 Cantonese prime sentences used in the main experiment and asked to translate each sentence into Mandarin; the other half were similarly asked to translate the 32 English prime sentences into Mandarin. Participants were equally accurate (i.e., produced the intended Mandarin verb) in Cantonese as in English (.87 vs. .85;  $t_s < 1.2$ ,  $p_s > .1$ ).

*Phonological similarity rating.* The Cantonese and Mandarin translation-equivalent verbs used in Experiment 2 were cognates and hence had related phonological forms (e.g., *maai-mai*; ‘buy’), whereas the English and Mandarin translation-equivalent verbs were non-cognates and hence did not have related phonological forms (e.g., *buy-mai*). Previous research found that complete phonological overlap of the head (here, the verb) between prime and target (e.g., *bat/bat*) may increase priming (Santesteban, Pickering & McLean, 2010). However, Cai et al. (2011) found that priming does not vary as a function of phonological

overlap of the dative verb. Following Cai et al. (2011), we had an additional 32 participants rate the Cantonese-Mandarin verb pairs using a 7-point Likert scale (1: *the pair of verbs sound very similar*; 7: *the pair of verbs sound very different*) (for results see <https://osf.io/znk37/>). If phonological similarity between translation-equivalent verbs affects the magnitude of priming, we should expect a larger priming effect for phonologically more similar pairs of translation equivalents.

*Procedure and scoring.* These were as in Experiment 1.

## Results

Table 2 presents the proportion of priming as a function of prime language. There was a significant intercept ( $\beta = 0.33$ ,  $SE = 0.09$ ,  $z = 3.83$ ,  $p < .001$ ), indicating that participants were more likely to use the syntactic structure of the prime sentence than the alternative structure. Separate analyses showed priming from Cantonese to Mandarin ( $\beta = 0.51$ ,  $SE = 0.11$ ,  $z = 4.55$ ,  $p < .001$ ), but not from English to Mandarin ( $\beta = 0.13$ ,  $SE = 0.09$ ,  $z = 1.51$ ,  $p = .132$ ). Critically, prime language was a significant predictor ( $\beta = 0.19$ ,  $SE = 0.07$ ,  $z = 2.88$ ,  $p = .004$ ): Priming was stronger from Cantonese to Mandarin (62%) than from English to Mandarin (53%) (see Fig. 2). Further analysis showed no effects of block order (all  $ps > .1$ ).

Pearson correlation tests showed that translation accuracy did not correlate significantly with priming magnitude for either Cantonese ( $r = 0.107$ ,  $p = .558$ ) or English ( $r = 0.153$ ,  $p = .402$ ), and phonological similarity did not correlate significantly with priming magnitude for Cantonese ( $r = -.081$ ,  $p = .661$ ). These results

suggest that neither translation accuracy nor phonological similarity (consistent with Cai et al., 2011) was related to the magnitude of priming.

To further explore whether repetition of verb meaning modulated the effect of language similarity on structural priming, we conducted a comparison between Experiments 1 and 2. We first established that participants in the two experiments did not differ in language proficiency or age of acquisition (all  $ps > .1$ ). We then analyzed the combined data, treating experiment (Experiment 1 vs. 2) and prime language (Cantonese vs. English) as fixed factors, and participant and item as random factors. The results showed that participants tended to reuse the syntactic structure in the prime (as indicated by the significant intercept,  $\beta = 0.25$ ,  $SE = 0.05$ ,  $z = 5.25$ ,  $p < .001$ ) and this tendency was similar between the two experiments (as indicated by the non-significant main effect of experiment,  $\beta = 0.13$ ,  $SE = 0.10$ ,  $z = 1.41$ ,  $p = .159$ ). The priming effect was larger from Cantonese than English to Mandarin (as indicated by the significant main effect of prime language,  $\beta = 0.09$ ,  $SE = 0.05$ ,  $z = 1.96$ ,  $p = .050$ ). More critically, the effect of prime language was modulated by experiment ( $\beta = 0.19$ ,  $SE = 0.09$ ,  $z = 2.11$ ,  $p = .035$ ): Priming from Cantonese to Mandarin was larger when verb meaning was repeated (i.e. Experiment 2; 62% primed responses) than when it was not (i.e. Experiment 1; 54% primed responses) ( $\beta = 0.32$ ,  $SE = 0.13$ ,  $z = 2.53$ ,  $p = .012$ ), whereas priming from English to Mandarin was similar whether or not meaning verb was repeated (55% and 53% respectively in Experiments 1 and 2;  $\beta = -0.06$ ,  $SE = 0.13$ ,  $z = -0.44$ ,  $p = .660$ ) (see Fig. 2).

## **Discussion**

In contrast to Experiment 1, which showed no difference in Cantonese-Mandarin and English-Mandarin priming, Experiment 2 showed stronger Cantonese-Mandarin than English-Mandarin priming when primes and targets involved translation-equivalent verbs. Both Experiments 1 and 2 used written target preambles, hence the language similarity effect in Experiment 2 cannot be attributed to shared overall orthography. To exclude the possibility that it might arise from shared orthography of the prime/target verb, we carried out a further experiment that was identical to Experiment 2 except that participants heard the target preambles. The experiment also aims to further examine whether there is priming from English to Mandarin when the prime and target have translation equivalent verbs.

## **Experiment 3**

### **Method**

*Participants.* 68 further participants (19 male, 18–24 years; mean 20.41) from the same population as Experiment 1 were paid 25 RMB. Four participants were discarded (three due to a technical issue in response recording and one withdrawing early in the experiment). As in previous experiments, participants reported acquiring Mandarin before Cantonese, and Cantonese before English, and speaking Mandarin with greater overall proficiency than Cantonese and English, (*t*-tests, all *ps* < .05). Their self-ratings for Cantonese and English proficiency did not differ (all *ps* > .1), except for listening comprehension (*t* = 3.12, *p* = .003) and age of acquisition (*t* = -5.12,

$p < .001$ ).

*Materials, Procedure, and Scoring.* These were identical to Experiment 2, except that target preambles were presented auditorily.

*Results and discussion.* Analysis was as in Experiment 1. The results replicated Experiment 2: There was a significant intercept ( $\beta = .24$ ,  $SE = .04$ ,  $z = 5.44$ ,  $p < .001$ ), indicating that participants were more likely to use the syntactic structure used in the prime sentence than the alternative structure, both from Cantonese to Mandarin ( $\beta = .35$ ,  $SE = .06$ ,  $z = 5.47$ ,  $p < .001$ ), and from English to Mandarin ( $\beta = .14$ ,  $SE = .06$ ,  $z = 2.20$ ,  $p = .028$ ). The priming from English to Mandarin contrasted that null finding in Experiment 2. Critically, prime language was a significant predictor ( $\beta = .21$ ,  $SE = .09$ ,  $z = 2.33$ ,  $p = .02$ ): The priming effect was stronger from Cantonese to Mandarin (59%) than from English to Mandarin (53%) (see Fig. 2). As we used spoken preamble in this experiment, such a finding suggests the stronger priming effect from Cantonese than from English did not arise from shared orthography between the prime and target verbs. There were no effects of block order (all  $ps > .1$ ).

## **General Discussion**

Our results show that multilinguals share abstract syntactic information between languages that are not similar overall as well as between languages that are similar overall, and critically that such sharing is unaffected by language similarity. In contrast, language similarity does affect the way in which multilinguals link words to syntax. Multilingual speakers showed equivalent cross-linguistic structural priming

between similar and dissimilar language pairs when the prime and target did not involve translation-equivalent verbs (Experiment 1) but stronger priming between similar languages than between dissimilar languages when the prime and target involved translation-equivalent verbs (Experiments 2 and 3).

Cai et al. (2011) showed that bilinguals of closely related languages have shared syntax for equivalent constructions (e.g., datives) but separate lemma representations for cognate translation equivalents. The current results extend these findings by further showing that syntactic representations are no more integrated but translation equivalents are more associated between more than less similar languages.

Importantly, this pattern of effects cannot be explained in terms of differences between similar/dissimilar language pairs with respect to participants' proficiency (e.g., Runnqvist et al., 2013), or the frequency of alternative structures (Benolet & Hartsuiker, 2010). Our participants' (self-rated) proficiency in Cantonese and English did not differ on the vast majority of measures (13 out of 15 measures), and moreover there was no relationship between proficiency and magnitude of priming on the two measures where there were significant differences (listening comprehension, Experiments 1 and 3). Additionally, although there were differences between languages in the frequency of the target structures (DO structures are more frequent in Mandarin and English than in Cantonese; see Branigan et al., 2000; Cai et al., 2011), previous research suggests that such differences would cause consistently stronger priming following Cantonese DO primes than English DO primes in all three

experiments (Bernolet & Hartsuiker, 2010), contrary to our findings.

We interpret our findings within an integrated lexical-syntactic account that was developed to explain representation and processing in monolinguals (Pickering & Branigan, 1998) but generalized to multilinguals (Hartsuiker et al., 2004). It is based on the model of lexical access developed by Levelt, Roelofs, and Meyer (1999). It assumes that language-independent concepts are linked to language-specific lemmas (i.e., the syntactic component of a lexical entry), which are in turn linked to language-independent combinatorial nodes (capturing the syntactic structures in which the word can occur). Priming is the result of residual activation of combinatorial nodes; stronger priming when a verb (or its translation-equivalent) is repeated results from residual activation of combinatorial nodes and strengthening of the link between a lemma and a combinatorial node (note that our account makes no assumptions about the dynamics of initial activation, e.g., serial vs. cascading activation).

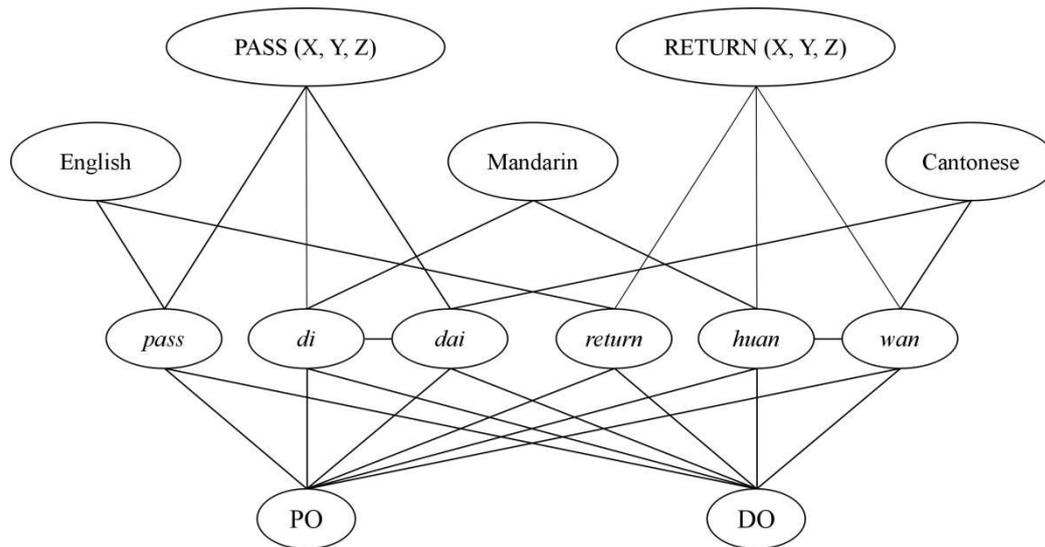
In our study, the equivalence of Cantonese-Mandarin and English-Mandarin abstract priming (i.e., in the absence of translation equivalent verbs) implies that where possible, speakers represent combinatorial potential (through combinatorial nodes) in a way that is fully integrated between languages, independent of the similarity of those languages to each other (and indeed to other languages that they may know – for example, Hakka and Chaoshanese have many similarities to Mandarin and Cantonese). It also implies that the PO and DO constructions in English

are linguistically equivalent to those in Cantonese and Mandarin, a finding that supports the use of structural priming as a method of investigating linguistic representation (Branigan & Pickering, 2017). In contrast, the stronger Cantonese-Mandarin than English-Mandarin priming found with verb repetition implies that the representation of lemmas – which link lexical and syntactic information – is influenced by language similarity. That is, the lemmas for cognate verbs in Cantonese and Mandarin (e.g., *dai* and *di*) are more closely linked than the lemmas for semantically equivalent non-cognates in English and Mandarin (e.g., *pass* and *di*). This closer link cannot take place via the conceptual representation, which is the same in all languages (i.e., when the meaning is the same).

Instead, we propose that speakers of Cantonese and Mandarin develop a direct link between the lemmas for cognate verbs (see Fig. 3), as a result of the extended confluence of activation at different levels between languages. For example, every time such speakers hear the Cantonese verb *dai* used in a particular structure, they activate not only the Cantonese-specific lemma *dai* but also the concept  $PASS(X,Y,Z)$ , the relevant combinatorial node (e.g., DO), and, importantly, the orthography, all of which are shared with Mandarin. (The same would be true when they heard a cognate verb in another related language such as Hakka or Chaoshanese). Because of this sharing, activation of  $PASS(X,Y,Z)$ , the PO node, and the orthography in turn activate the Mandarin lemma *di*. (In addition, the Cantonese phonology /*dai*/ overlaps with the Mandarin phonology /*di*/, and this overlap may also support shared activation, though

we note that priming was not enhanced for more versus less phonologically similar verbs in Mandarin and Cantonese.) Co-activation of the Mandarin and Cantonese lemmas via orthographic (and potentially phonological) links thus leads to the development of a link between them through Hebbian learning (Munakata & Pfaffly, 2004). As Hebbian learning occurs over linguistic input or output (e.g., comprehension or production of *dai* activates *di* via shared phonology/orthography), it is likely that the link gradually develops over the course of bilingual development, with weaker links for child than adult bilinguals. Such a learning mechanism is also consistent with the proposal that the cognate advantage in word recognition is a frequency effect in disguise (Lalor & Kirsner, 2001; Strijkers et al., 2010); for instance, the perception or production of *dai* would also lead to activation of *di*, hence increasing the frequency of *di*. In contrast, between English and Mandarin, there is no equivalent link between translation equivalents – there is shared conceptual and combinatorial activation, but no shared activation of orthography (or phonology), and hence no link develops between the *pass* and the *di* lemmas.

Fig. 3. Model of lexico-syntactic representations in Mandarin-Cantonese-English multilinguals.



This means that multilinguals' syntactic representations are not more integrated for similar languages than for dissimilar languages. Overall, Cantonese and Mandarin lemmas are connected to the same combinatorial nodes in the same way as English and Mandarin lemmas (where the three languages have equivalent structures). But there is a closer integration of representations that connect lexical and syntactic information for similar than for dissimilar languages. In conclusion, language similarity affects how multilinguals represent their languages – but with respect to the syntactic characteristics of individual words rather than the language as a whole.

### **Author contribution**

All authors developed the study concept and contributed to the study design. Testing and data collection were performed by X. Chen. Z. Cai and J. Huang performed the data analysis. H.P. Branigan and M.J. Pickering drafted the manuscript and all authors provided critical revisions. All authors approved the final version of the manuscript for submission.

### **Acknowledgments**

This research was supported by the Royal Society of Edinburgh, National Science Foundation of China, National Social Science Foundation of China, National Natural Science Foundation of Guangdong Province, Key Institute of Humanities and Social Sciences of China (grants 15AZD048, 2014A030311016, 31500880, 16JJD880025), and Economic and Social Research Council (ES/L010224/2).

## References

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278. doi: <http://dx.doi.org/10.1016/j.jml.2012.11.001>
- Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, 114(3), 455–61. <https://doi.org/10.1016/j.cognition.2009.11.005>
- Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355-387. doi: [https://doi.org/10.1016/0010-0285\(86\)90004-6](https://doi.org/10.1016/0010-0285(86)90004-6)
- Branigan, H.P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, 40. doi: 10.1017/S0140525X17001212
- Cai, Z. G., Pickering, M.J., Yan, H., & Branigan, H.P. (2011). Lexical and syntactic representations in closely related languages: Evidence from Cantonese—Mandarin bilinguals. *Journal of Memory and Language*, 65(4), 431-445. doi: <https://doi.org/10.1016/j.jml.2011.05.003>
- De Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics*, 13, 1–24.
- Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493-498.

- Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*.  
Cambridge MA: Harvard University Press.
- Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science, 15*(6), 409-414. doi: DOI 10.1111/j.0956-7976.2004.00693.x
- Kantola, L., & van Gompel, R. P. G. (2011). Between- and within-language priming is the same: Evidence for shared bilingual syntactic representations. *Memory & Cognition, 39*(2), 276-290. doi: 10.3758/s13421-010-0016-5
- Lalor, E., & Kirsner, K. (2001). The representation of “false cognates” in the bilingual lexicon. *Psychonomic Bulletin & Review, 8*, 552–559.
- Language Assignments to 3/3 LDPs. (2015). Retrieved from <https://www.state.gov/documents/organization/247092.pdf>
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-38.
- Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics, 41*(5), 791-824. doi: DOI 10.1515/ling.2003.026
- Munakata, Y., & Pfaffly, J. (2004). Hebbian learning and development. *Developmental Science, 7*(2), 141-148. doi: DOI 10.1111/j.1467-7687.2004.00331.x
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and*

*Language*, 39(4), 633-651. doi: DOI 10.1006/jmla.1998.2592

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review.

*Psychological Bulletin*, 134(3), 427-459. doi: 10.1037/0033-2909.134.3.427

Runnqvist, E., Gollan, T. H., Costa, A., & Ferreira, V. S. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition*, 129(2), 256–263.

<https://doi.org/10.1016/J.COGNITION.2013.07.008>

Santesteban, M., Pickering, M. J., & Mclean, J. F. (2010). Lexical and phonological effects on syntactic processing: Evidence from syntactic priming. *Journal of Memory and Language*, 63(3), 347-366. DOI: 10.1016/j.jml.2010.07.001

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M.J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56(2), 153-171. doi:

<https://doi.org/10.1016/j.jml.2006.10.002>

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, 20, 912–928.

Van Gompel, R.P. G., & Arai, M. (2017). Structural priming in bilinguals.

*Bilingualism: Language and Cognition*, 1-8. doi:

10.1017/S1366728917000542

Wagenmakers, E. (2007). A practical solution to the pervasive problems of p values.

