Running head: Cantonese-English loanword priming

Full title: **No significant loanword priming advantage in Cantonese-English bilinguals**

Stephen Politzer-Ahles, Wing Ki Ng, Li Chong Shih

The Hong Kong Polytechnic University

Address for correspondence:

Stephen Politzer-Ahles

The Hong Kong Polytechnic University

Hong Kong

sjpolit@polyu.edu.hk

E-mail addresses of other authors:

Wing Ki Ng: kate.wingki.ng@connect.polyu.hk

Li Chong Shih: rachel.shih@connect.polyu.hk

Abstract

Bilingual speakers can process a word more quickly if it is primed by its translation equivalent than if it is primed by an unrelated word. This translation priming effect is particularly large when the words are cognates or loanwords. This pattern, however, has only been tested in bilinguals for whom both languages are written with phonology-based writing systems. In this high-power study we test whether there is a similar magnification of the translation priming effect in loanwords when the two languages use different kinds of writing systems. Using data from 200 Cantonese-English bilinguals, we find significant translation priming but no advantage for loanword pairs relative to non-loanword pairs. This suggests that the priming advantage reported for loanwords and cognates in many languages is not cross-linguistically general, but may depend on certain properties shared by the two languages in question, such as the writing system or the nature of the relation between the loanwords and their translation equivalents.

Keywords: masked priming, translation priming, loanwords, Chinese, orthography

Introduction

One of the major areas of research interest in the psycholinguistics of bilingualism is how the respective vocabularies of a bilingual's languages are organized—i.e, whether words from the two languages are represented separately, or are connected/integrated within the speaker's mental lexicon. Much research in this area has used the *translation priming* paradigm (see, e.g, Wen & van Heuven, 2017, for review). Priming refers to the phenomenon whereby people can more quickly or more easily process a stimulus after having encountered a related stimulus. In a typical priming experiment, for example, English might perform a task in which they must press a button as quickly as possible to decide if a word is a real word or not. In such a situation, when they are shown *DOCTOR* they will tend to press the button more quickly if they have just seen *nurse* than if they have just seen *table*. A priming effect provides evidence that the processing of one stimulus (in this example, the word *nurse*) activates representations that are related to it (e.g, the word *doctor* or its associated concept), facilitating their subsequent processing. Translation priming is the facilitation obtained when a word is preceded by its translation equivalent in another language. For example, a French-English bilingual will tend to respond faster to *DOG* if they have just seen *chien* (which means "dog" in French) than if they have just seen *oiseau* (which means "bird" in French).

Importantly, this translation priming effect (i.e, the difference in reaction time between *oiseau*$_{\text{"bird"}}$…*DOG* and *chien*$_{\text{"dog"}}$…*DOG*) varies as a function of the relation between these translation equivalents. Some translation equivalents, like *chien*$_{\text{"dog"}}$ and *dog*, have no obvious relationship other than their shared meaning. Other translation equivalents, however, have form similarities. For instance, French *acteur*$_{\text{"actor"}}$ and English *actor* both mean the same thing, and have the same spelling and similar pronunciations; they also both derive from the Latin word *actor*.[i] Pairs like these are cognates. Similarly, French *shampooing* and English *shampoo* have the same meaning and have similar spellings and pronunciations; these words do not come from one common ancestor between the languages, but rather the word was borrowed from English into French, and thus is a loanword.

Crucially for our present purposes, translation priming effects tend to be larger for word pairs that are cognates or loanwords than for word pairs that are not; that is to say, French-English word pairs like *acteur*…*ACTOR* or *shampooing*…*SHAMPOO* tend to elicit more translation priming than pairs like *chien*$_{\text{"dog"}}$…*DOG*. This pattern, known as the *cognate priming* *advantage*, has been observed in numerous priming experiments (see Table 1, also available as a comma-separated spreadsheet at https://osf.io/kubgc/). There is some debate over whether this cognate priming advantage occurs because cognate or loanword pairs actually share a lexical representation in a way that non-cognate, non-loanword pairs do not, or whether it is just because the phonological relatedness between the words provides some extra facilitation over and above that provided by their translation equivalence (see Nakayama et al, 2014, and Voga & Grainger, 2007, among others; this is consistent with the assumptions of BIA+ [Djikstra & van Heuven, 2002], a prominent model of bilingual lexical representation, which assumes that a bilingual's lexicons are integrated and that cognates or loanwords are not represented in a qualitatively different way than unrelated translation equivalents are but simply have more form overlap). Regardless of which account is true, neither assumes that the advantage should have a different locus for cognates than for loanwords. Given that the difference between cognates and loanwords is primarily diachronic, we are not aware of any reason to predict that these should have different kinds of representations the mind of a typical speaker, who likely is naïve to etymology. Therefore, here we treat cognates and loanwords as equivalent here, in keeping with the rest of the priming literature on this topic, which generally makes no distinction between cognates and loanwords; this means that when we refer to "loanword priming" or the "loanword priming advantage" we are referring to the same thing that is often called "cognate priming" or the "cognate priming advantage" in other studies.

<Insert table 1 about here>

The loanword priming advantage (or cognate priming advantage) has been found in pairs of languages written with different scripts, such as English and Hebrew (Gollan, Frost, & Forster, 1997), English and Korean (Kim & Davis, 2003), French and Greek (Voga & Grainger, 2007), English and Urdu (Khan, 2012), and English and Japanese (Nakayama et al, 2013); see Table 1 for a detailed summary (also available as a comma-separated spreadsheet at https://osf.io/kubgc/). Such results are important, because the way translation equivalents are represented in the minds of bilinguals whose languages share similar scripts (like French and English) is not necessarily representative of the way translation equivalents are represented in the minds of all types of bilinguals. Examining loanword priming in languages written with different scripts provides a good test for whether the loanword priming advantage is simply because loanwords have strong form similarity (e.g, visual/graphic similarity) across languages, or because their representations in the mental lexicon are somehow more closely connected across languages than representations of non-loanword translation equivalents are. The research done so far, however, has just scratched the surface of the possible relationships between bilinguals' lexical representation of translation equivalents and their orthographic knowledge. Specifically, almost all extant studies we are aware of that tested the loanword priming advantage examined bilinguals whose languages are both written with sound-based writing systems. Whether their characters mostly represent phones (as in alphabets like Basque, Dutch, English, French, Greek, Korean, and Spanish), consonants (as in Hebrew and Urdu, both impure abjads), or morae (as in Japanese kana), these writing systems roughly represent the sound structure of a language. (Note, however, that the correspondence between sounds and characters in these writing systems is not always perfectly one-to-one, especially in highly irregular writing systems like English.) Few studies have examined whether there is a loanword priming advantage in bilinguals whose languages are written with entirely different classes of writing system—as opposed to writing systems that are different on the surface but are both phonological in nature.

To address this question, we conducted a high-power[ii] study examining loanword and non-loanword translation priming in bilingual speakers of Cantonese and English. Cantonese is generally written with Chinese characters,[iii] which are sometimes referred to as a "logographic" script. In a true logographic writing system, each character represents a word. This is not true of Chinese; the vast majority of characters in Chinese represent one or more morphemes,[iv] and many characters are not used alone as stand-alone words in modern Chinese languages. Nonetheless, regardless of how exactly Chinese is categorized, it stands in stark contrast to the sound-based writing systems used in previous research on the loanword priming advantage. While Chinese characters do contain some phonetic information—the majority of characters are made up of sub-character components called radicals, which often roughly correlate with the pronunciation or meaning of a character—the representation of pronunciation information in Chinese characters is substantially more opaque than in the other writing systems mentioned above. Our aim in the present study, then, was to test whether a loanword priming advantage would also be observed in Cantonese-English bilinguals, given that their writing systems are even less similar than those compared in previous studies. A recent study on Mandarin-English bilinguals (Zhang, Wu, Zhou, & Meng, in press) observed such a cognate advantage, as did a similar study by Qi (2011); the present study tests for this in a substantially larger sample and with rigorous statistical methods, using Cantonese-English bilinguals.

We considered two predictions for this study. On the one hand, it was possible that the loanword priming advantage would be replicated in this study, given that Cantonese-English loanwords have phonological similarity just like, e.g, Japanese-English loanwords do. On the other hand, there could be an advantage for *non*-loanwords. All non-loanwords used in this study were multimorphemic, as most content nouns in Cantonese are compounds consisting of two morphemes. Their constituent morphemes are often related to the meaning of the whole word. For example, 醫院 ("hospital") is made up of the morphemes 醫 (*ji*^{1}, "medicine"; pronunciations are given in the Jyutping system of Romanization and the superscript numbers indicate tone categories) and 院 (*jun*^{6}, "court"/"institution"). Thus, the processing of such words may be facilitated by activation from their constituent parts, and this activation may likewise increase the activation of their translation equivalents. Loanwords, on the other hand, tend to be monomorphemic words, made up of characters that are mostly just used is in phonetic loanwords and which are no longer used meaningfully on their own in modern Chinese (or at least have no transparent relationship to the meaning of the whole word). For example, 芭菲 ("parfait") is made up of 芭 (*baa*^{1}) and 菲 (*fei*^{1}), which are characters whose use is mainly limited to loanwords, and which do not contribute any meaning to the meaning of the two-syllable word. These loanwords, then, may fail to be semantically facilitated by activation of their constituent parts in the same way as typical words are. For this reason, even though we predicted a loanword priming advantage for Cantonese-English bilinguals, it is nonetheless possible that there could be a loanword priming *disadvantage*.

The study we conducted was pre-registered. Pre-registration is a technique to increase reliability of and confidence in experimental findings (Nosek, Ebersole, DeHaven, & Mellor, 2017; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012) gaining popularity in psychological sciences recently. It is meant to address the issue of flexibility in data analysis. There are many choices that can be made during data analysis, and it has long been known that trying out many different analyses, and selectively reporting just the ones that yielded the most attractive results, increases the likelihood of publishing implausible or un-replicable findings (Simmons, Nelson, & Simonsohn, 2011). This is the case even when the analysis choices look justifiable after the fact (Gelman & Loken, 2013). The best way to address this issue is to make all analysis decisions before having looked at the results, and keep a record of these decisions. This record is a pre-registration: a plan for how the data will be analysed. Even if a pre-registration is not followed exactly, the pre-registration record makes it easier to distinguish between analyses that were planned ahead of time (and which can thus be considered *confirmatory*, offering evidence for a hypothesis) versus analyses or analysis steps that were only added after the data had been seen (and which thus must be considered *exploratory*, offering new hypotheses that still need to be tested in future research). For the present study, given that the results could turn out to be very sensitive to analysis choices such as how to identify outliers, which covariates to include, etc, we pre-registered the analysis plan in order to avoid trying out different analyses and settling on one that supports our preferred hypothesis.

Methods

All methods for this study were pre-registered with the Open Science Foundation at https://osf.io/dm2yh/. Where our analysis deviates from the pre-registered plan, we indicate so in the text. All relevant materials, data, and analysis code are provided at https://osf.io/kubgc/.

Participants

200 volunteers (53 men, 147 women) from Hong Kong participated in the experiment. They were from 18 to 35 years old (mean: 25) and spoke Cantonese as a native language.[v] They all had high English proficiency, based on standardized test scores—we solicited volunteers with International English Language Testing System (IELTS) scores of 6 or above, Hong Kong Certificate of Education Examination levels of D or above, or Hong Kong Diploma of Secondary Education levels of 3 or above (for detailed demographic information, see https://osf.io/kubgc/; the latter two tests mentioned above are commonly used English proficiency certifications in Hong Kong); no English proficiency test was administered because we had no *a priori* hypotheses about interactions with proficiency. On average they had 19 years of English education (range: 12-30), and the age at which they begun to have English instruction, according to self-report, was age 6 on average (range: 0-19). All participants provided their informed

consent and were compensated with cash for their participation. Half of the participants completed the experiment with Cantonese primes and English targets, and half with English primes and Cantonese targets; this was manipulated between participants rather than within participants since this comparison was not the primary aim of the study, and manipulating it between participants allowed each individual participant to contribute more trials per cell. Several participants might be considered early, balanced bilinguals, and classifying them as having had "L2 primes" or "L1 primes" may not be accurate; these participants are indicated in the "Notes" column of the demographic information sheet at https://osf.io/kubgc/.

Materials

Fifty-four Cantonese loanwords of English origin were used, along with their English translation equivalents (e.g, 巴士 [*ba*^{1} *si*^{2}, "bus"] and *bus*), as well as fifty-four Cantonese non-loanwords along with their English translation equivalents (e.g, 紅酒 [*hung*^{4} *zau*^{2}, "wine"] and *wine*).[vi] The Cantonese words were all 2-3 syllables/characters in length. Each Cantonese word was paired with a related prime (its English translation equivalent) as well as another English word unrelated in form or meaning to the Cantonese word. Likewise, each English word was paired with a related Cantonese prime (its translation equivalent) as well as an unrelated Cantonese word. The unrelated primes always had the same number of characters as the corresponding related primes. Example stimulus sets are shown in Table 2; the full stimulus list is available at https://osf.io/kubgc/.

<Insert table 2 about here>

We suspected that loanwords and non-loanwords might differ in frequency (how commonly they are used) and concreteness—for example, many Cantonese loanwords from English are words for foods, and technology. Therefore, we collected these measures for each word so that we could later ensure that any differences in priming effects are not due merely to differences in frequency or concreteness. Because there is no large comprehensive Cantonese corpus with information on word frequency and concreteness, and because it would be difficult to compare data across different corpora which may have very different properties,[vii] we instead estimated these lexical properties using a survey. We distributed surveys in which volunteers rated each word (English target, Cantonese target, unrelated English prime, and unrelated Cantonese prime; the related primes were the same as the targets in the other languages) in terms of their concreteness (on a 6-point scale from "very concrete" to "very abstract") and estimated age of acquisition (on a 7-level scale including 0-3 years, 3-6, 6-9, 9-12, 12-15, 15-18, and 18 or above). Age of acquisition is known to be closely positively correlated with word frequency (Juhasz & Rayner, 2007). Each word was rated by five participants,[viii] and the ratings were averaged across participants to yield a value for each word (excluding responses of "don't know", for age of acquisition, or "cannot judge", for concreteness). We note that five ratings per item is a fairly small sample size, so these estimates of abstractness and age of acquisition are fairly rough.

108 Cantonese nonword fillers and 108 English nonword fillers were also included. The Cantonese nonwords were made by taking two existing characters and combining them in a way that does not yield an existing compound word, e.g, 照思. The English nonwords were phonotactically legal pseudowords chosen from the ARC Nonword Database (Rastle et al, 2002). Each nonword had a phonologically unrelated prime in the opposite language (Cantonese nonword targets had English primes, and English nonword targets had Cantonese primes).

Priming direction (whether the experiment included L1 primes with L2 targets, or L2 primes with L1 targets) was manipulated between participants, as the effect of priming direction was not of primary interest in this study and manipulating it within participants would have caused each individual to have fewer trials per condition. Thus, each participant saw 54 loanword targets (half with related primes and half with unrelated primes) and 54 non-loanword targets (half with related primes and half with unrelated primes) in the same language, along with 108 nonword targets in that same language. For each direction of priming, the 108 critical trials with word targets were arranged into two lists in a Latin square design, such that each target was only seen once per participant, with either a related target or an unrelated target. Within each language direction, the nonword fillers were the same across both lists.

Procedure

Participants were tested one by one in a quiet room. Stimulus presentation and data logging was carried out using the DMDX software package (Forster & Forster, 2003). Each trial began with a forward mask ("######") presented for 500 ms, followed by a prime that remained on screen for 67ms.[ix] The target word was then presented immediately after the prime and then remained on screen until either the participant responded, or 4000ms elapsed. English words (whether prime or target) were presented in lowercase. Because some prime words take up more horizontal space than their corresponding targets (e.g. prime *saxophone* with target 色士風, or prime 模特兒 with target *model*), all targets were flanked by brackets when being presented (e.g. ">>>色士風<<<" or ">>>model<<<"), in order to make sure that the primes were being completely masked (methodology based on Nakayama et al, 2014).[x] Participants were instructed to judge whether each target (e.g. "色士風") was a real word or not as quickly and accurately as possible. They indicated their answer by pressing the right shift key (representing "yes") or left shift key (representing "no").

The experiment began with 20 practice trials to help participants acclimate to the experiment. After this, the 216 experimental trials (108 critical trials and 108 fillers) were presented. The experimental items were pseudorandomly arranged using the DMDX default procedure, with a scramble block size of 16; this means that the 216 items were divided into sets of 16 trials, these sets were arranged in a random order, and then the trials within each set were arranged in a random order. Each set included 2 trials of each condition (loanword with related target, loanword with unrelated target, non-loanword with related target, and non-loanword with unrelated target) and 8 nonword filler trials. Trials and sets were presented continuously one after the other, except for two breaks given during the experiment, dividing the experiment into three blocks (plus practice).

Analysis

Trials with incorrect responses were excluded from subsequent analyses. Subsequently, any trials with reaction times more than 1.5 interquartile intervals away from the median response time for that participant or item were marked as outliers and excluded from further analysis. Statistical analyses were conducted with linear mixed effect models (Baayen, Davidson, & Bates, 2008), with maximal random effects justified by the design (Barr, Levy, Scheepers, & Tily, 2013), carried out in R (R Core Team, 2016). Models included nuisance covariates for age of acquisition and concreteness of the primes and targets;[xi] for details, see the analysis code at https://osf.io/kubgc/. Reaction time was log-transformed (this transform yielded the least skewed model residuals) and z-scored to reduce the likelihood of model convergence failures. All predictors were centered so that their model coefficients would be interpretable in the presence of interactions; the continuous covariates were z-scored and the categorical predictors (prime relatedness, target loanword status, and priming direction) were deviation-coded. Two effects were of interest: the interaction between prime relatedness (related vs. unrelated) and target loanword status (loanword vs. non-loanword), and the three-way interaction between these two factors and priming direction (L1-L2 vs. L2-L1). If loanwords show a larger priming effect than non-loanwords, this would elicit a two-way interaction (we used a one-tailed test, since this is a directional prediction). If that interaction is limited to just one priming direction, that would elicit a three-way interaction (we used a two-tailed test, since we had no prediction about which priming direction would show a bigger loanword priming advantage). Since evidence for a loanword priming advantage could have come from either of two different model coefficients, we set the alpha level to .025.

Results

Prime awareness

While no explicit prime awareness test was administered, we did experience that many participants with primes in their native language (L1-L2 direction) reported after the experiment that they could see the primes. This was not reported in the L2-L1 directions. This is consistent with our experience running the script (two of the authors are Cantonese native speakers and could see the Cantonese primes but not the English primes, whereas one author is an English native speaker and could see the English primes but not the Cantonese primes).

Manipulation checks

Figure 1 shows the priming effects (reaction time for unrelated trials minus reaction time for related trials) for each participant, for both loanwords and non-loanwords. It is clear that there is robust priming: the majority of participants show slower reaction times for unrelated than related trials. The statistical model (reproducible code for all analyses is available on https://osf.io/kubgc/) revealed a 44-ms main effect of prime relatedness which was highly significant (t=14.41), confirming that the experiment was able to elicit masked translation priming effects. It is also clear that the priming effect for L1-L2 priming was substantially larger than the priming effect for L2-L1 priming, consistent with previous reports (e.g, Wen & van Heuven, 2017); in the statistical model, the priming effect for L1-L2 priming was 73 ms larger than that for L2-L1 priming and this difference was highly significant (t=10.391).

<Insert Figure 1 about here>

Pre-registered analysis

While the effects illustrated above serve as a manipulation check to confirm that the experiment was sensitive to priming effects, the interactions involving loanword status are the only ones relevant to the research question. Figure 2 shows the loanword priming advantage (the priming effect for loanwords minus that for non-loanwords) for each participant. If there is a loanword priming advantage, this value should be positive. It is apparent from visual inspection that if there is any loanword priming it is not very robust across participants: almost as many participants have a loanword priming *disadvantage* as a loanword priming advantage. Accordingly, the small interaction between relatedness and loanword status—in the statistical model, the priming effect for loanwords is 2 ms larger than for non-loanwords—was not significant (t=0.24; this corresponds roughly to one-tailed p=.405, depending on the method used to estimate p-values). Figures 1 and 2 suggest that there may be some trend towards a loanword advantage in the L1-L2 direction but very little such trend in the L2-L1 direction; there was not, however, a significant interaction between prime relatedness, loanword status, and priming direction to support such a conclusion (t=0.57, roughly two-tailed p=.566). While this result is subject to the same limitations as all non-significant inferential statistical tests (Altman & Bland, 1995), we note that (1) this experiment likely had larger power than most others in this field (see Footnote 1); (2) the study had sufficient power to strongly detect priming in general, even L2-L1 priming which is often difficult to detect (Wen & van Heuven, 2017), so it is not likely that the experiment was too insensitive to detect loanword advantage; and (3) as shown in Figure 2, there is hardly even a trend towards a loanword advantage in the L1-L2 direction, and no trend at all in the L2-L1 direction, whereas if the null result were due only to insufficient power we would expect to see clear but non-significant trends.

<Insert Figure 2 about here>

Exploratory analyses

An anonymous reviewer pointed out that the model used in this pre-registered analysis may be affected by multicollinearity (see, e.g, Baayen, Feldman, & Schreuder, 2006). We checked the variance inflation factors for this model, and indeed, while most were between 2 and 5, the variance inflation factors (based on estimated condition R^2 [Nakagawa & Schielzeth, 2013]) for the target properties were in the millions, indicative of serious multicollinearity for these factors (at least if they are included in the random effects structure). We thus ran an exploratory analysis in attempt to reduce this concern. Firstly, we removed prime abstractness and prime age of acquisiton from the model, as we did not expect them to have substantial impact on reaction times (given that the reaction times were to the target, not to the prime). Secondly, given that abstractness and age of acquisition are correlated (later-acquired words tend to be more abstract), we ran a principal component analysis on these two factors in attempt to reduce them to a single dimension. The first principal component accounted for 82% of the variance in target abstractness and age of acquisition, so we ran a new model using only each target's weight on this component, instead of each target's abstractness and age of acquisition. The results were qualitatively similar to those of the pre-registered analysis. There was a highly significant main effect of prime relatedness (t=18.27) and a highly significant interaction between relatedness and priming direction (t=11.25), but not a significant interaction between relatedness and loanword status (t=1.03) nor a significant three-way interaction between relatedness, loanword status, and direction (t=1.32). (The interation between relatedness and the principal component was fairly strong and negative, t=-2.07; since high values on this component were associated with late-acquired, abstract words, and low values with early-acquired, concrete words, this interaction suggests that the priming effect was stronger in early-acquired, concrete words.)

The patterns shown in Figures 1 and 2 do not exactly reflect the results of the pre-registered statistical model, since the model included numerous covariates that are not shown in these figures (see Analysis). For ease of exposition, we also fitted a model without covariates (see analysis code at https://osf.io/kubgc/), to more directly correspond to the patterns shown in the figures. In this model the interaction between prime relatedness, loanword status, and priming direction is still not significant (t=1.16), but the crucial two-way interaction between prime relatedness and loanword status is marginal significant (t=1.32, roughly one-tailed p=.094 uncorrected). While this might be taken as evidence that there was a loanword priming advantage, we believe such a conclusion would be premature. First of all, this analysis was not pre-registered and is thus exploratory, and subject to all the caveats of unplanned analyses (e.g, Simmons, Nelson, & Simonsohn, 2011). More importantly, this analysis does not control for lexical differences between loanwords and non-loanwords. While loanwords and non-loanwords had similar ages of acquisition (see stimulus properties at https://osf.io/kubgc/; Cantonese loanword and non-loanword targets each had average ages of acquisition of 3.2 on the 7-point scale, and English loanword and non-loanword targets had 3.7 and 3.8 respectively), loanwords were much more concrete (Cantonese loanword and non-loanword targets had average abstractness ratings of 1.9 and 2.4, respectively, on the 6-point abstractness scale; English targets had 1.8 and 2.3 respectively)—as mentioned above, loanwords tend to be names of foods, vehicles, articles of clothing, technological artifacts, etc. Thus, we cannot rule out the possibility that apparent differences in priming effect size between loanwords and non-loanwords may have instead been due to these differences in lexical properties; indeed, the fact that including those properties as covariates diminishes the loanword advantage suggests that that was the case.

Discussion

In a pre-registered study with fairly high power (compared to typical sample sizes in experiments on this topic) on L1 Cantonese – L2 English speakers, we found very little evidence that Cantonese-English loanwords elicit a larger masked priming effect than non-loanwords. The failure to observe such a loanword priming advantage is very unlikely to be due to a lack of power, given that the study was sufficiently powered to detect even very small priming effect sizes (see, e.g, the small but significant L2-L1 priming in Figure 1). The other aspects of the experiment replicated commonly found patterns, e.g, that masked translation primes elicit facilitation and that this facilitation is smaller for L2 primes to L1 targets than for L1 primes to L2 targets (Wen & van Heuven, 2017); the fact that these patterns were observed serves as a manipulation check to increase confidence that the experiment worked as intended and elicited typical translation priming effects. What sets this experiment apart from the other literature on loanword priming is that there are almost no other studies using a lexical decision task with masked primes that failed to find a loanword priming advantage (see Table 1). The vast majority of similar studies found a priming advantage for loanwords or cognates. This suggests that some aspect of Cantonese-English bilingualism has caused the results to be different than the results of other language pairs previously studied.

One possibility is that the different natures of the languages' two writing systems is the cause for the difference. Cantonese is typically written with a script that is mostly morpheme-based, whereas English is typically written with an alphabetic script. While loanword or cognate priming advantages have been documented in many pairs of languages with different writing systems (see Table 1), all pairs previously studied are pairs in which both languages use phonologically-based writing systems, broadly defined (see Introduction). The present study is one of the only studies, and the largest yet, to examine cross-language priming in bilinguals who speak one language written with a phonologically-based writing system and one with a non-phonologically-based writing system. A cross-language loanword priming advantage for phonologically-based writing systems but not for non-phonologically-based writing systems does not seem to be directly predicted by the phonological account of cognate/loanword priming effects (Voga & Grainger, 2007), which just predicts that cognates or loanwords should elicit larger priming effects as long as their phonological form is similar across the two languages. If the effect is in fact limited to languages in which the phonological form can be more or less transparently read off the orthographic representation, this would warrant an update to that account. As this is the first study to suggest such a cross-language difference, this conclusion is tentative and needs further validation before such a change would be justified. If the explanation outlined above is correct, then one might predict that, for example, less transparent orthographies (like English words with irregular spelling) would elicit less loanword/cognate priming advantage than more transparent orthographies (ones with fewer exceptions) like Spanish or Korean.

Another possible explanation for the lack of significant loanword priming advantage in the present study relates to the morphological makeup of the words. As described in the Introduction, the Cantonese non-loanwords used in the study are mostly multi-morphemic compound words, where the meaning of each constituent character contributes to the meaning of the word. On the other hand, the loanwords are mostly monomorphemic, and the constituent characters are only there for phonological purposes and do not contribute meaning. It is possible (although by no means has it been empirically demonstrated before, as far as we know) that compound primes or targets can elicit greater priming effects than simple ones, given that a bimorphemic compound word may have three parts contributing to eventual activation of its lexical representation (the compound word itself, and each of its constituent parts along with their meanings) whereas the simple word only has one. If this is the case, then whatever loanword priming advantage is present in the study may have been counteracted by a comparable loanword priming disadvantage due to the monomorphemic nature of the loanwords. This is a post-hoc speculation that requires further testing, both to confirm the hypothesis that multi-morphemic words elicit greater translation priming than monomorphemic ones, and to confirm that this can cancel out a loanword priming advantage. It would also be valuable to test whether a loanword priming advantage can be observed in Cantonese-English bilinguals when using loanwords that actually are multi-morphemic. Phono-semantic loanwords may qualify as such. For instance, the Cantonese word for *laser* is 鐳射, made up of the meaningful syllables 鐳 (*leoi*^{4}, "radium") and 射 (*se*^{6}, "to shoot"); while the word is borrowed from and sounds similar to the English source, it is also made up of two morphemes that have at least some relation to the meaning of "laser". Mandarin also has such loanwords, e.g, 跑酷 ("parkour", literally *pao*^{3} "to run" and *ku*^{4} "cool") and 黑客 ("hacker", literally *hei*^{1} "wicked" and *ke*^{4} "visitor"). If there are enough such words in the language, they could be a useful additional test case for loanword priming between English and Cantonese or Mandarin.

Related to the above point, it is possible that loanwords and non-loanwords are processed via different routes. As mentioned above and in the introduction, non-loanwords have semantic cues to their meaning (in both their constituent characters, and in the constituent characters' sub-character radicals), whereas loanwords often do not, and might instead need to be cued more by the phonological forms represented by the characters. These processing routes may differ in speed or efficiency, given that non-loanwords are more common. Another thing that could make loanword processing difficult is that representations of a loanword's constituent characters might compete with the representation of the word itself. For instance, 芭菲 ("parfait") includes the character 芭, which in this word is only used because of its pronunciation, but which originally referred to a kind of herb; for speakers who are aware of that meaning, it may interfere with accessing the meaning of the whole word 芭菲. These possibilities are speculative and would need empirical demonstration, but in any case they cannot account for the present results; even if loanword processing is, for some reason, slower or more effortful than non-loanword processing across the board, that does not explain why *priming* for loanwords would be similar to priming for non-loanwords (i.e, any two words with vastly different overall reaction times could still have the exact same priming effect sizes).

The present study did take phonetic similarity between loanwords and their translation equivalents into account. As pointed out by an anonymous reviewer, some of the loanwords are phonetically fairly similar to their translation equivalents (e.g, 巴士 [*ba*^{1} *si*^{2}, "bus"]), whereas some are not so similar (e.g, 忌廉 [*gei*^{6} *lim*^{1}, "cream"]). It is has been previously demonstrated that the loanword priming advantage may be reduced or absent when the phonetic similarity is low (e.g, Nakayama et al, 2014). Thus, another potential explanation for the lack of a loanword priming advantage in the present study could be that the phonetic similarity between the English and Cantonese words was too low in too many items. Whether phonetic similarity can account for the apparent lack of loanword advantage in Cantonese-English bilinguals is an empirical question that remains to be addressed. This is a promising route for study, but we note that it may be difficult to operationalize how much more or less similar Cantonese-English loanwords are than loanwords in other pairs of languages. Cantonese is a tone language, and it is not clear how much weight tones should be given in evaluation of phonological similarity even within a language (e.g, Yao & Sharma, 2017), much less across languages (while there are somewhat regular associations between Chinese tones and English phonology, they are not one-to-one; see, e.g, Jian, 2017). It also still remains unclear, in a variety of applications, what the unit of phonological similarity evaluation should be; e.g, while the number of overlapping phonemes is often used as a measure of phonological similarity, this could just as meaningfully be measured by overlapping features (e.g, pseudowords *paf* and *baf* may be more phonologically similar than pseudowords *paf* and *zaf*, because even though both pairs differ in exactly one phoneme, the former pair differs in just one feature—voicing—whereas the latter differs in three features—voicing, place, and manner). This problem is compounded for comparisons across different languages, where the phoneme and feature inventories may not be the same. And other aspects beyond segments, such as syllable structure and numbers of syllables, may also influence similarity (e.g, comparing the pseudoword pair *tilk* and *tilf* with the pair *tilk* and *tila*, even though they each differ in one phoneme, might the second pair be considered more different since it also differs in number of syllables?). One way to avoid this challenge is to rely instead on subjective ratings of phonological similarity (e.g, Nakayama et al, 2014), which presumably are influenced by many of these factors. However, it is not clear that subjective ratings of phonological similarity between, say, Cantonese and English, would be directly comparable to subjective ratings of phonological similarity between, say, Spanish and English; since a rating scale is subjective, the raters for these two pairs might unconsciously apply different standards. All that is to say, showing that Cantonese-English loanwords are less phonetically similar to their translation equivalents than other language pairs are would be challenging. An alternative way to get evidence for this possibility may be to show, within just Cantonese-English loanwords, that there is indeed a loanword priming advantage when looking only at the highest-similarity items.

The results of the present study are inconsistent with those of Zhang and colleagues (in press) and Qi (2011), who all observed a cognate priming advantage in Chinese-English bilinguals.[xii] Those studies used substantially smaller samples of participants than the present study (69 in the study by Zhang and colleagues and 41 in the study by Qi, compared to 200 in the present study), although the sample sizes of items were similar (24 trials per condition per participant in the study by Zhang and colleagues and 38 in the study by Qi, compared to 27 in the present study. Larger-scale replication directly comparing Mandarin and Cantonese speakers is probably necessary to confirm whether there is a robust difference between these populations (rather than the difference between these studies being due to Type 1 or Type 2 error on the part of one or more studies). It is not clear why Cantonese-English bilinguals in Hong Kong would have a different effect of loanword status on priming than Mandarin-English bilinguals in China (Zhang et al, in press) and Singapore (Qi, 2011) would; this is an open question for future study.

As noted above, the primes L1-L2 direction in the present experiment may not have been subliminal for many participants. However, if anything this strengthens the present findings: the L1-L2 direction is the direction in which there was a slight, albeit non-significant, trend towards loanword advantage (Figure 2), whereas in the L2-L1 direction, where the primes were indeed subliminal, there is no trace of a trend towards loanword advantage. We also note that loanword priming advantage has been observed at least one study with a longer prime-target stimulus onset asynchrony (and thus possibly more visible primes) than the present study (Ferré, Sánchez-Casas, Comesaña, & Demestre, 2017). Therefore, we doubt that the visibility of the primes can account for the failure to observe a significant loanword priming advantage in the present study. The present study does show, however, that future studies using Cantonese-English loanword priming in a similar paradigm may consider using shorter stimulus onset asynchronies between the primes and targets.

Conclusion

The fact that the loanword priming advantage is not observed in Cantonese-English translation priming, unlike most other language pairs tested, challenges current understanding of the factors that modulate translation priming, and particularly the role of loanword status. The findings suggest that the pattern previously observed in many other languages may not be cross-linguistically general. Rather, it may be dependent on language-specific factors. While the present study cannot definitively determine what language-specific factors modulate the size of the loanword priming advantage, some possible candidates are the relationship between the writing systems of the two languages, or the morphological nature of the loanword process as opposed to native-language lexical items. Both of these topics would be valuable avenues for future research to elucidate the mechanisms that support translation priming effects, and bilingual lexical recognition in general.

Declarations

Availability of data and materials. The datasets generated during the current study, as well as the materials used, are available in the Open Science Foundation repository, https://osf.io/kubgc/

Competing interests. The authors declare that they have no competing interests.

Funding. This research was supported by grant G-UACX to SPA from the Departmental General Research Fund of the Department of Chinese and Bilingual Studies at the Hong Kong Polytechnic University. The funding body had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Author contributions. SPA contributed to the conception and design of the experiment, analysed the data, and wrote the manuscript. WKN and SLC contributed to the conception and design of the experiment, designed the materials, and collected the data. All authors read and approved the manuscript.

Acknowledgements. We would like to thank Lei Pan, Ming Chun Woo, and Ching Ying Yau for assistance with data collection. Exploratory data analysis benefitted from discussion with Si Chen. Any errors are our own.

References

Baayen, R, Davidson, D, and Bates, D. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390-412.

Baayen, R, Feldman, L, and Schreuder, R. 2006. Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 53, 496-512.

Barr, D, Levy, R, Scheepers, C, & Tily, H. 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Altman, D, & Bland, J. 1995. Statistics notes: absence of evidence is not evidence of absence. *BMJ, 311*, 485.

de Groot, A, & Nas, G. 1991. Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*, 90-123.

Davis, C, Sánchez-Casas, R, García-Albea, J, Guasch, M, Molera, M, & Ferré, P. 2010. Masked translation priming: varying language experience and word type with Spanish-English bilinguals. *Bilingualism: Language and Cognition, 13*, 137-155.

Dijkstra, T, & Van Heuven, W. 2002. The architecture of the bilingual word recognition system: from identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197.

Duñabeitia, J, Dimitropoulou, M, Morris, J, & Kiependaele, K. 2013. The role of form in morphological priming: evidence from bilinguals. *Language and Cognitive Processes, 28*, 967-987.

Duñabeitia, J, Perea, M, & Carreiras, M. 2010. Masked translation priming effects with highly proficient simultaneous bilinguals. *Experimental Psychology, 57*, 98-107.

Ferré, P, Sánchez-Casas, R, Comesaña, M, & Demestre, J. 2017. Masked translation priming with cognates and noncognates: is there an effect of words' concreteness? *Bilingualism: Language and Cognition, 20*, 770-782.

Forster, K, & Forster, J. 2003. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124.

Funahashi, T, & Yamana, H. 2010. Reliability verification of search engines' hit counts. *Proceedings of the 10th International Conference on Current Trends in Web Engineering.*

Gelman, A, & Loken, E. 2013. The garden of forking paths: why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. Unpublished manuscript.

Gollan, T, Forster, K, & Frost, R. 1997. Translation priming with different scripts: masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1122-1139.

Jian, L. 2017. Tonal adaptation of English loanwords in contemporary Mandarin. *Journal of Chinese Linguistics, 45*, 432-450.

Juhasz, B, & Rayner, K. 2007. The role of age of acquisition and word frequency in reading: evidence from eye fixation durations. *Visual Cognition, 13*, 846-863.

Khan, Q. 2013. *Lexical representation and processing in cross-script Urdu-English bilinguals: the case of frequency-balanced and frequency-unbalanced cognates and noncognates.* Ph.D. dissertation, University of Windsor.

Kim, J, & Davis, C. 2003. Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language, 49*, 484-499.

Leung, M, & Law, S. 2002. HKCAC: The Hong Kong Cantonese adult language corpus. *International Journal of Corpus Linguistics, 6*, 305-325.

Liberman, M. 2005. Questioning reality. *Language Log*. http://itre.cis.upenn.edu/~myl/languagelog/archives/001837.html

Loftus, G, & Masson, M. 1994. Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476-490.

Mair, V. 2011. Polysyllabic characters in Chinese writing. *Language Log*. http://languagelog.ldc.upenn.edu/nll/?p=3330. Accessed on 31 October 2017.

Nakayama, M, Sears, C, Hino, Y, & Lupker, S. 2013. Masked translation priming with Japanese-English bilinguals: interactions between cognate status, target frequency and L2 proficiency. *Journal of Cognitive Psychology, 25*, 949-981.

Nakayama, M, Verdonschot, R, Sears, C, & Lupker, S. 2014. The masked cognate translation priming effect for different-script bilinguals is modulated by the phonological similarity of cognate words: further support for the phonological account. *Journal of Cognitive Psychology, 26*, 714-724.

Nosek, B, Ebersole, C, DeHaven, A, & Mellor, D. 2017. The preregistration revolution. Preprint. https://osf.io/2dxu5/. doi: 10.17605/OSF.IO/2DXU5

Politzer-Ahles, S, & Fiorentino, R. 2013. The realization of scalar inferences: context sensitivity without processing cost. *PLoS ONE, 8*, e63943.

Qi, Y. 2011. *Cognate status and cross-script priming with Chinese-English bilinguals and English-Chinese bilinguals*. M.A. thesis, National University of Singapore.

R Core Team 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rastle, K, Harrington, J, & Coltheart, M. 2002. 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55A*, 1339-1362.

Sánchez-Casas, R, García-Albea, J, & Davis, C. 1992. Bilingual lexical processing: exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology, 4*, 293-310.

Simmons, J, Nelson, L, & Simonsohn, U. 2011. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359-1366.

Snell, J, Declerck, M, & Grainger, J. 2017. Parallel semantic processing in reading revisited: effects of translation equivalents in bilingual readers. *Language, Cognition and Neuroscience, 33*, 563-574.

Tse, C, Yap, M, Chan, C, Sze, W, Shaoul, C, & Lin, D. 2017. The Chinese Lexicon Project: A megastudy of lexical decision performance for 25,000+ traditional Chinese two-character compound words. *Behavior Research Methods, 49*, 1503-1519.

Voga, M, & Grainger, J. 2007. Cognate status and cross-script translation priming. *Memory & Cognition, 35*, 938-952.

Wagenmakers, E, Wetzels, R, Borsboom, D, van der Maas, H, & Kievit, R. 2012. An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*, 632-638.

Wen, Y, & Van Heuven, W. 2017. Non-cognate translation priming in masked priming lexical decision experiments: a meta-analysis. *Psychonomic Bulletin & Review, 24*, 879-886.

Yao, Y, & Sharma, B. 2017. What is in the neighborhood of a tonal syllable? Evidence from auditory lexical decision in Mandarin Chinese. *Proceedings of the Linguistic Society of America, 2*, 45:1-14.

Zhang, J, Wu, C, Zhou, T, & Meng, Y. in press. Cognate facilitation priming effect is modulated by writing system: evidence from Chinese-English bilinguals. *International Journal of Bilingualism*.

---

[i] *Etymonline* entry for "actor": http://www.etymonline.com/index.php?term=actor; *Wiktionary* entry for "actuer": https://en.wiktionary.org/w/index.php?title=acteur&oldid=47310844#French

[ii] It was not possible to calculate a specific estimate of the power for this study, as power depends the expected raw effect size and variance structure of the data. In a research design with crossed participants and items, the variance structure of the data is fairly complicated and thus a common way to calculate power is to conduct simulations, using the variance components from an existing dataset to simulate the new ones (for an example, see Politzer-Ahles & Fiorentino, 2013). In the present context, without such a previous dataset on which to model the expected variance structure, it would not have been feasible to calculate power estimates for any *a priori* raw effect sizes. Rather, we opted to simply collect a much larger sample than is usually used in similar experiments; thus, while we cannot know the exact power we had to detect any particular effect size, we can infer that our power to detect such an effect size is larger than studies with smaller samples, all else being equal.

[iii] We say "generally" because (1) some Cantonese morphemes, particularly sentence-final particles, do not have corresponding characters or are commonly written with Latin characters rather than their Chinese characters, particularly on computers or mobile devices; and (2) not all Cantonese speakers in the world are literate in Chinese characters. Note that Cantonese is typically written with traditional Chinese characters in Hong Kong and in many communities in the global Chinese diaspora outside of China, whereas it is typically written with simplified Chinese characters in the Cantonese-speaking regions of the People's Republic of China.

[iv] Exceptions include characters that represent multi-morpheme, polysyllabic words (Mair, 2011), and characters that represent sounds without any particular meaning (e.g, the characters making up the loanwords in this study).

[v] Several additional volunteers participated and were not included in this analysis: one participant whose data were lost, one who misunderstood the instructions and judged whether the targets were related to primes rather than whether the targets were words (this data exclusion was not pre-registered), and 22 who did not meet the pre-registered English proficiency criteria (18 of whom had scores in tests other than the three listed in our pre-registration; and 4 of whom had scores below the minimum cutoff on those tests listed in our pre-registration). The data for all these participants (except the lost dataset) are included at https://osf.io/kubgc/ and can be analyzed using the code provided there (simply by removing or editing the line that excludes them from the eventual analysis).

[vi] An additional six loanword pairs and six non-loanword pairs were created, but were removed from the experiment after stimulus norming (see below) because their age of acquisition was substantially higher than the others or because people were unfamiliar with the written forms of some Cantonese loanwords used mainly in speech. The actual experiment did not include these twelve items; they were instead among the practice trials.

[vii] A recent Cantonese corpus based on spoken language exists (Leung & Law, 2002), but is fairly small (about 170,000 syllables). Another recent corpus (Tse, Yap, Chan, Sze, Shaoul, & Lin, 2017) estimated frequencies for Cantonese characters using the Hong Kong traditional Chinese character database on Google, but we are not convinced that this accurately reflects spoken word frequency, given substantial register differences between written (formal) and spoken (informal) Cantonese; furthermore, Google hit counts are only estimates and are often substantially inaccurate, particularly for search queries with a large number of hits (e.g, Funahashi & Yamana, 2010; Liberman, 2005). Finally, regardless of the merits of any individual corpus, our planned analysis includes covariates related to both prime and target properties in the same model; since primes and targets are in different languages, this would necessitate using English and Chinese frequency counts obtained from different corpora with different kinds of texts. For these reasons, we opted not to use corpus estimates of frequency, but instead to use subjective ratings. Finally, it is important to note that frequency is only an estimate of some underlying, un-observable construct (an individual's experience with a given word over their lifetime), and self-rated age of acquisition, which

---

we used instead of frequency, is also an estimate of this same construct, and is associated with many of the same variables that frequency is (see, e.g, Juhasz & Rayner, 2007)

[viii] Each participant rated one-third of the Cantonese words or one-third of the English words, and provided ratings for both variables. The rating was administered via Google Surveys. No fillers or catch items were included.

[ix] 60ms was specified in the DMDX script, but the experiment was run on two laptop computers which each had monitor refresh rates of about 16.7ms. Thus, DMX would have actually presented the stimuli for 4 refresh cycles, or 67ms.

[x] A potential concern with this method is that parafoveal information from flankers can impact processing of the target words in reaction time studies (e.g, Snell, Declerk, & Grainger, 2018). We believe this is not a confound for the present study because previous evidence for parafoveal-on-foveal effects is based on designs in which the flankers are real words, whereas in our experiment the flankers are semantically neutral, meaningless syllables, and in any case they are the same across all conditions.

[xi] Other potentially relevant nuisance covariates include those related to visual complexity of the stimuli, such as length and, for Chinese characters, stroke count. We did not include these in this model because it is not clear that they could be compared across English and Chinese targets, where things like stroke count and number of characters have substantially different meanings. However, these variables are included in the stimulus files online, and would be straightforwardly included in models that analyze Chinese and English targets separately.

[xii] The advantage Zhang and colleagues (in press) observed was significant for individuals tested in the L2-L1 direction but not for those tested in the L1-L2 direction, but the advantages in these two cases were numerically almost identical. The authors do not report either direct comparisons or pooled analyses across the two directions. Therefore, we conclude that the most reasonable conclusion is that they did not observe significantly loanword priming advantages across priming directions.

Table 1. *Summary of extant studies on loanword/cognate priming advantage, including the present study. The values for the priming effect sizes for the present study are taken from the statistical model, and thus are adjusted for the presence of covariates.*

Experiment	$N_{participants}$	Prime language	Target language	Script relationship	Task	Prime-Target SOA (ms)	Cognate/loanword priming effect (ms)	Non-cognate / non-loanword priming effect (ms)	Cognate/loanword advantage (ms)
De Groot & Nas (1991, Experiment 3)	68	Dutch (L1)	English (L2)	Same alphabet	Lexical decision	240 (unmasked)	68	113	-45
De Groot & Nas (1991, Experiment 3)	68	Dutch (L1)	English (L2)	Same alphabet	Lexical decision	60 (masked)	48	35	13
De Groot & Nas (1991, Experiment 4) collapsing across case conditions	76	Dutch (L1)	English (L2)	Same alphabet	Lexical decision	60 (masked)	70	31	39
Sánchez-Casas et al. (1992, Experiment 1)	21	English (mostly L2)	Spanish (L1)	Same alphabet	Semantic categorization	60 (masked)	32	-8	40
Gollan et al. (1997, Experiment 1)	40	Hebrew (L1)	English (L2)	One alphabet and one impure abjad	Lexical decision	50 (masked)	53	36	17

Gollan et al. (1997, Experiment 2)	30	English (L1)	Hebrew (L2)	One alphabet and one impure abjad	Lexical decision	50 (masked)	142	52	90
Gollan et al. (1997, Experiment 3)	40	English (L2)	Hebrew (L1)	One alphabet and one impure abjad	Lexical decision	50 (masked)	9	9	0
Gollan et al. (1997, Experiment 4)	30	Hebrew (L2)	English (L1)	One alphabet and one impure abjad	Lexical decision	50 (masked)	4	-4	8
Kim & Davis (2003, Experiment 1)	25	Korean (L1)	English (L2)	Different alphabets	Lexical decision	50 (masked)	34	40	-6
Kim & Davis (2003, Experiment 2)	18	Korean (L1)	English (L2)	Different alphabets	Naming	50 (masked)	28	8	20
Kim & Davis (2003, Experiment 4)	16	Korean (L1)	English (L2)	Different alphabets	Semantic categorization	50 (masked)	52	58	-6
Voga & Grainger (2007, Experiment 2), collapsing across phonological overlap	30	Greek (L1)	French (L2)	Different alphabets	Lexical decision	50 (masked)	50	23	27
Voga & Grainger (2007, Experiment 3)	30	Greek (L1)	French (L2)	Different alphabets	Lexical decision	50 (masked)	48	22	26
Duñabeitia et al. (2010) collapsing	32	Basque and Spanish (L1)	Basque and Spanish (L1)	Same alphabet	Lexical decision	47 (masked)	53	18	35

across language directions									
Davis et al. (2010, Experiment 1) collapsing across speaker group and language direction	84	English and Spanish (L1 and L2)	English and Spanish (L1 and L2)	Same alphabet	Lexical decision	57 (masked)	22	-10	32
Davis et al. (2010, Experiment 2) collapsing across language direction	21	Spanish and English (L1 and L2)	Spanish and English (L1 and L2)	Same alphabet	Lexical decision	57 (masked)	29	5	24
Qi (2011, Experiments 1 and 2)	21	Mandarin (L1)	English (L2)	One alphabet and one logosyllabary	Lexical decision and naming	50 (masked)	59	37	22
Qi (2011, Experiments 3 and 4)	21	English (dominant L1)	Mandarin (less dominant L1)	One alphabet and one logosyllabary	Lexical decision and naming	50 (masked)	75	20	55
Khan (2012, Experiment 2b) collapsing across frequency and language direction	29	Urdu and English (L1 and early dominant L2)	Urdu and English (L1 and early dominant L2)	One alphabet and one impure abjad	Lexical decision	30 and 50 (masked)	-8	14	-22

Khan (2012, Experiment 2c)	25	English (early dominant L2)	Urdu (L1)	One alphabet and one impure abjad	Lexical decision	50 (masked)	65	34	30
Nakayama et al. (2013, Experiment 1) collapsing across frequency and proficiency	66	Japanese (L1)	English (L2)	One alphabet and one other system (logographic for non-loanwords, moraic for loanwords)	Lexical decision	50 (masked)	81	59	22
Nakayama et al. (2013, Experiment 2B)	32	English (L2)	Japanese (L1)	One alphabet and one other system (logographic for non-loanwords, moraic for loanwords)	Lexical decision	50 (masked)	15	-1	16
Duñabeitia et al. (2013, Experiment 1)	44	Spanish (L1)	English (L2)	Same alphabet	Lexical decision	50 (masked)	25	-8	33
Ferré et al. (2016, Experiment 1) collapsing across direction and concreteness	32	Spanish and English (L1 and L2)	Spanish and English (L1 and L2)	Same alphabet	Lexical decision	50 (masked)	32	15	17

Ferré et al. (2016, Experiment 2) collapsing across direction and concreteness	38	Spanish and English (L1 and L2)	Spanish and English (L1 and L2)	Same alphabet	Lexical decision	100 (masked)	42	20	21
Zhang et al. (in press, Experiment 1)	35	Mandarin (L1)	English (L2)	One alphabet and one logosyllabary	Lexical decision	60 (masked)	68	42	26
Zhang et al. (in press, Experiment 2)	34	English (L2)	Mandarin (L1)	One alphabet and one logosyllabary	Lexical decision	60 (masked)	26	-2	28
Present study, L1-to-L2 direction	100	Cantonese (L1)	English (L2)	One alphabet and one logosyllabary	Lexical decision	67 (masked)	80	75	5
Present study, L2-to-L1 direction	100	English (L2)	Cantonese (L1)	One alphabet and one logosyllabary	Lexical decision	67 (masked)	9	11	-2

Table 2. *Example stimuli. English translations of the Cantonese words are shown in parentheses. "L2-L1" refers to the experiment using second-language (English) primes and first-language (Cantonese) targets. "L1-L2" refers to the experiment using first-language (Cantonese) primes and second-language (English) targets.*

Condition	Target	Related Prime	Unrelated Prime
L2-L1 loanwords	巴士 (bus)	*bus*	*cat*
L2-L1 non-loanwords	紅酒 (wine)	*wine*	*fire*
L1-L2 loanwords	*bus*	巴士 (bus)	學生 (student)
L1-L2 non-loanwords	*wine*	紅酒 (wine)	窗簾 (curtain)

Figure captions

Figure 1. Priming effects for each participant. Observations are horizontally jittered to reduce visual overlap. Thick horizontal black lines show condition means, and error bars indicate one-tailed 95% confidence intervals of the priming effects estimated from a mixed-effects model without lexical covariates. Each confidence interval shows the range of points that that priming effect is not significantly different from; as such, they can be compared against zero to show that each priming effect is significant in of itself, but they cannot be compared against each other to see if one priming effect is different than another, as these comparisons include repeated-measures data (see, e.g., Loftus & Masson, 1994).

Figure 2. Loanword priming advantage. Loanword priming advantage (priming effect size for loanwords minus priming effect size for non-loanwords) for each participant. Observations are horizontally jittered to reduce visual overlap. Thick horizontal black lines show means, and error bars indicate one-tailed 95% confidence intervals of the loanword priming advantage estimated from a mixed-effects model without lexical covariates.