Introduction

Bilingualism is the norm in most of the world's population, and many people use one or more second languages in their everyday life. Thus, understanding how the mental architecture supports language use and language comprehension in bilingual speakers, and how this may or may not differ from more-studied monolingual and first-language-using populations, is very important if our theory of the psychology of language is to be ecologically valid for describing how language actually works in real-world situations. The present research project will how words are comprehended in Cantonese-English bilinguals, specifically focusing on loanwords, which are a less studied aspect of the bilingual speaker's lexicon.

Commented [PS[1]]: Here I raise the big-picture idea. My actual proposal will be much more specific than this, but this big-picture idea helps show why this kind of experiment could be important.

Commented [PS[2]]: Here I make the question a bit more specific

A previous eye-tracking experiment (Van Assche, Duyck, Hartsuiker, and Diependaele, 2009) has suggested that skilled bilingual readers read a word faster (in a sentence context) if it is a cognate or loanword in two or more languages they speak; for instance, a Dutch-English bilingual might read the Dutch word *schip* (meaning "ship", with spelling and pronunciation very similar to its English counterpart) faster than the Dutch word *lade* (meaning "drawer", with no spelling or pronunciation relationship to its English counterpart). This suggests that such words are connected, across their languages, within the mental lexicon.

Commented [PS[3]]: Here I briefly describe the state of current knowledge (what we know about this topic so far). This also serves to boil down the big abstract question into a much more specific issue (whether or not people read loanwords faster than non-loanwords).

This study, however, has some limitations. Firstly, the conclusions are based on a between-items comparison (i.e., faster reading times for *schip* than for *lade*), so it is not possible to exclusively attribute the reading time difference to cognate status. These words differ from each other in many ways, not just in their status as cognates or non-cognates, so it is possible that some other un-measured variable is confounded with cognate status. For example, if the cognate words were more common or differed in any other way from the non-cognate words, then the apparent cognate effect may just have been an epiphenomenon of this other variable). On the other hand, work from other paradigms often includes a control group; e.g., Costa, Caramazza, and Sebastian-Galles (2000) found that in a picture naming task, words with cognates can be said more quickly than words without cognates, but only in a group of bilingual speakers and not in a monolingual control group; similarly, Bice and Kroll (2015) also conducted an ERP experiment on cognate lexical decision with both bilinguals and monolingual controls.

An additional open question left by this study is that of whether or not the results will extend to bilinguals whose languages have very different writing systems. Dutch and English use roughly the same alphabet, and therefore Dutch-English cognates have substantial overlap in orthography. If it is true that readers read cognate words faster than non-cognate words, we don't yet know if this only happens when the two languages have substantial spelling overlap, or if this will also happen with language pairs with no spelling overlap.

Commented [PS[4]]: Here I identify the knowledge gap. In this proposal I actually am pointing out two gaps, but for your proposal you only need one.

The present study

The present study will adapt the design of Van Assche and colleagues (2009) to attempt to address these questions. Just like their experiment, this experiment will us eye-tracking to measure how quickly bilinguals read loanwords and non-loanwords in sentence contexts. To address the second question, I will use Cantonese-English bilinguals. Cantonese has many loanwords from English (e.g., 巴士, from English *bus*) which share meaning and pronunciation across the language but have no overlap in spelling. If the cognate reading advantage described by Van Assche and colleagues (2009) does not depend on orthographic overlap, then L1 English speakers who are fluent in Cantonese should read English words faster if they have Cantonese loanwords than if they do not (e.g., when reading a sentence, these speakers should read *bus* faster than they read *car*, since *bus* corresponds to the loanwords 巴士 whereas *car* corresponds to the non-loanword translation 汽車).

Secondly, to address the issue of between-items confounds, this study will use a control group (non-bilingual speakers). If words like *bus*, which have Cantonese loanwords, are read faster than words like *car*, which do not, then if this effect is because of their loanword status then this effect should only occur in bilinguals. Monolinguals should not show a difference in reading times to these words, since monolinguals don't know that *bus* has a Cantonese loanword equivalent. On the other hand, if this reading time difference is due to other confounds (e.g., if words like *bus* are higher-frequency than words like *car*), then the reading time difference will occur in both monolinguals and bilinguals.

However, this research has not been extended to bilinguals who use languages with different writing systems, like Cantonese and English; furthermore, extant studies investigated this question have confounds which limit the conclusions that can be made from their results. The present project will update this paradigm to test whether loanwords really are read faster than non-loanwords, and if this pattern occurs even for languages with no spelling overlap, like Cantonese and English.

Methods

All the analyses for this project will be pre-registered on the Open Science Foundation platform (http://osf.io) and all stimuli, analysis code, and data will be made publicly available there.

Participants

Approximately 180 volunteers without history of language impairment will participate in the experiment. The goal is to have at least 80 participants who are native English speakers with no Cantonese proficiency (beyond everyday phrases) and at least 80 who are native English speakers with advanced Cantonese proficiency; a slightly larger sample will be recruited to account for expected participant attrition—i.e., participants not attending well to the task, etc.). Power analysis is not feasible because I do not have *a priori* estimates of the expected

Commented [PS[5]]: Here I describe the proposed experiment, how it addresses the knowledge gaps, and what the predictions are

Commented [PS[6]]: Here I identify the knowledge gap. In this proposal I actually am pointing out two gaps, but for your proposal you only need one.

Commented [PS[7]]: This is something I have begun doing recently in my experiments. To learn about what pre-registration is, see https://cogtales.wordpress.com/2016/04/01/preregistration-its-actually-a-really-good-idea/. To learn about why sharing data publicly is important, see http://sometimesimwrong.typepad.com/wrong/2017/03/looking-under-the-hood.html and https://medium.com/@richarddmorey/science-deserves-a-gift-this-christmas-d4c911f806e4.

It's not required that you do this in your proposal. But I believe it's good practice.

Commented [PS[8]]: Here I describe how many volunteers I will recruit, and what kind variance structure in the data, thus I am instead using the more conservative alternative to power analysis, which is to take the typically sample size for current psycholinguistic eye-tracking studies (often around 20-40 participants per cell) and double it, based on widespread knowledge in the statistics community that current experiments are underpowered. Because of the specific nature of the sample, many volunteers will be recruited from outside the university community (e.g., professionals working in Hong Kong on short-term contracts and not learning the local language, and graduates of international schools with high Cantonese proficiency); the sample will likely be more diverse than typical samples in psycholinguistic experiments which are predominantly made up of university students.

Commented [PS[9]]: Here I justify why I am choosing this number of participants

Design

Volunteers will participate in an experiment in which they read English sentences with critical words that do or do not have a Cantonese loanword equivalent, as shown in (1) and (2) below:[1]

1) **With Cantonese loanword equivalent**: Ramesh | saw | a bus | on the | other side | of the street.
2) **Without Cantonese loanword equivalent**: Ramesh | saw | a car | on the | other side | of the street.

The translation of *bus* in Cantonese is a loanword (巴士), whereas the translation of *car* is not (汽車). Sentences will be constructed so that they do not generate a strong prediction towards either word. The design closely follows Van Assche et al. (2009), the only difference is that this study includes a control group of English speakers who do not know Cantonese, whereas that study did not include a corresponding control group. (The other difference is that these words are loanwords rather than cognates. True cognates are words that are similar between languages because of shared genetic background; e.g., English *state*, French *état*, and Spanish *estado* all derive from Latin *status*. Van Assche et al., 2009, used true cognate words. Cantonese and English have no true cognates because they do not have a common ancestor; translation equivalents in these languages which have similar pronunciations, like *bus* and 巴士 or *toast* and 多士 are instead loanwords/borrowings. We have no *a priori* reason, however, to predict that loanwords would behave differently than cognates in speakers' brains, because psycholinguistic theories about their representation are focused on the surface similarity of these words rather than on their etymologies [and indeed, many proficient native speakers probably do not know the etymologies of words in their language] and many studies [e.g., Nakayama et al., 2014] ignore this distinction

Commented [PS[10]]: Here I describe the experiment design in detail, explaining exactly what conditions will be compared, and giving examples.

[1] The "|" marks indicate how the sentence is divided up into regions for the eye movement analysis. They won't actually be shown in the experiment.

and assume that loanwords and cognates should behave the same way synchronically in speakers' minds.)

Materials

Following Van Assche and colleagues (2009), the stimuli will be 40 pairs of English words with and without Cantonese loanwords (e.g., *bus* and *car*) in 40 respective sentence frames. Each of the 40 sentence frames will fit about equally well with its two respective words (as in examples 1-2 above; goodness of fit will be confirmed with a sentence completion norming survey before the eye-tracking experiment is conducted). The reason for choosing to use 40 of each sentence type is 1) to match what was done by Van Assche and colleagues (2009); and 2) to keep the experiment to a manageable length to participants do not feel too fatigued.

Commented [PS[11]: Here I explain how many sentences will be used, give details about how they will be made (i.e., specifying that I will do a survey beforehand to make sure that all sentences have critical words that fit well in the context), and justifying why I choose this number of sentences

Procedure

The stimuli will be arranged into two lists following a Latin Square design, such that each sentence appears in only one condition (with-loanword or without-loanword) per participant, but every participant sees sentences from each condition [across different sentences] and every sentence appears in both conditions [across different participants].

The experiment proper will be conducted with an Eyelink 1000 system in the SLS lab hosted by the Department of Chinese and Bilingual Studies in QR602. At the beginning of the experiment, participants will perform 9-point calibration so that the eye tracker accurately measures where their eyes are looking; 9-point calibration will be repeated throughout the experiment any time the tracking appears inaccurate. Trials will be presented in a fully random order. At the beginning of each trial, participants will first fixate on a dot on the left-hand side of the screen for drift correction; after this, the sentence will appear for participants to read. There will be a gray box at the upper right-hand corner of the screen, which participants can look at once they finish the sentence in order to tell the experiment that they are done reading it (this method is often considered superior to allowing participants to end the sentence by pressing a button, because it prevents them from being able to end the sentence before they finish reading it). After 25% of the sentences, the participant will be shown a comprehension question with two possible answers (e.g., "Where did Ramesh see the car?", with choices "across the street" or "in the garage"), which they can respond to using buttons on the keyboard.

Commented [PS[12]: Most of this is specific detail based on previous eye-tracking experiments I have done. If you don't know all the details about the machine, you can just say something like "we will use the same procedures described in [some paper]". The most important part here is the part explaining what the participant's task is (i.e., saying that the participant will respond to a comprehension question after some sentences), and the part explaining the order trials will be presented (a random order) and the Latin Square design.

Data analysis

Eye movement data will be preprocessed according to the procedure described by Van Assche and colleagues (2009): outliers (defined as fixation times more than 2 standard deviations away from that participant's mean) will be removed, as will trials where the critical word not fixated or where the first fixation on the critical word was not progressive.[2]

I will analyze three eye movement measures, based on what was analyzed by Van Assche and colleagues (2009). Analysis will be done at the critical word itself (e.g., *a bus* or *a car*), and exploratory analysis will also be done on the following region in case there are spillover effects:

- *First fixation duration*: the duration of the first fixation on the region
- *First pass time* (Van Assche and colleagues refer to this as "gaze duration"): The sum of all fixations within a region from when the region was first entered until when the region was exited in either direction
- *Go-past time* (Van Assche and colleagues refer to this as "regression path duration"): The sum of all fixations (including fixations in previous regions) from when the region was first entered until when the region is exited to the right (i.e., until a fixation at a later region is made)

Commented [PS[13]: Here I describe the specific data analysis – the important thing is how outliers will be handled, and what reading time measures I will look at (on what words in the sentence)

Reading times will be statistically evaluated with linear mixed-effects models (Baayen, Davidson, & Bates, 2008) with fixed nuisance covariates for possible confounding variables (frequency, letter count, and cloze probability of the critical words) as well as critical fixed predictors of loanword status (loanword or non-loanword) and speaker group (English-Cantonese bilinguals or English monolinguals). First-pass times and go-past times will be log-transformed.[3]

Commented [PS[14]: This is a lot of details about statistics; if you don't know statistics you don't need to include this, you can just say something more simple (e.g., that you will compare the mean reading time in one condition to the mean reading time in the other condition).

Predictions

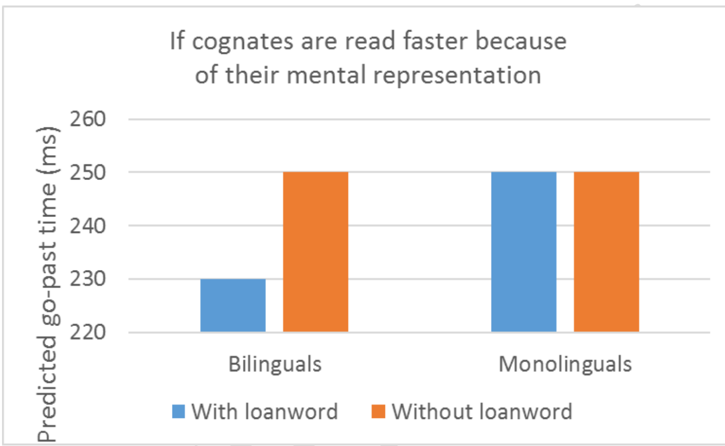
If cognates really are read more quickly, as claimed by Van Assche and colleagues (2009), then in at least one reading measure (first fixation time, first pass time, or

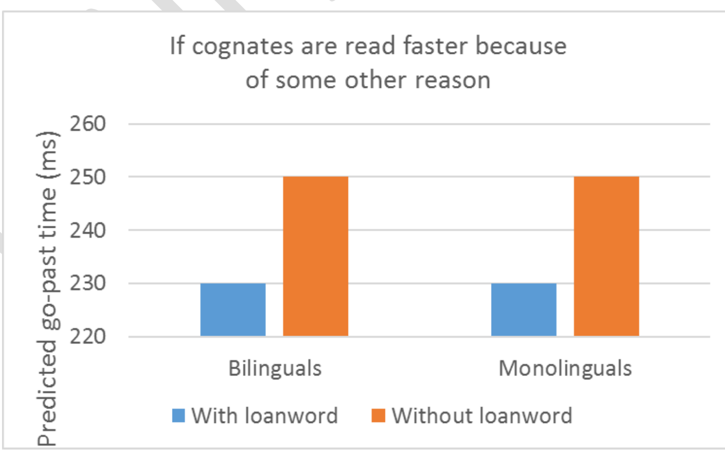
[2] This is done in order to replicate their analysis, to make the present results more comparable to theirs. As an exploratory analysis, I will also analyze the data using what I consider a more ideal analysis plan, with data cleaning in four steps based on the defaults in the Eyelink Data Viewer program. First, fixations of 80 ms or shorter will be merged into a neighboring fixation of greater than 80 ms within 0.5 degrees horizontally (if both the preceding and following fixation are longer than 80 ms, the short fixation will be merged to the longer of the two); second, the same process will be repeated with a duration threshold of 40 ms and a distance threshold of 1.25 degrees; third, in interest areas that had at least three fixations of 140 ms or shorter and none of longer than 140 ms, the short fixations will be merged into one; and last, remaining fixations shorter than 140 ms or longer than 800 ms will be deleted. After this, outliers for each reading time measure will be identified based on visual inspection (following Baayen, 2008) rather than on a standard deviation threshold. Trials with incorrect answers to comprehension questions will still not be removed, since comprehension questions will not have been presented with all trials.

[3] This is done to replicate what was done by Van Assche and colleagues; in reality, it does not always make sense to log-transform. For an exploratory analysis, I will instead do what I usually do: test the skewness of the residuals from models computed on raw, square-root transformed, natural log (log base *e*) transformed, common log (log base 10) transformed, and reflected-residual transformed data, and then only analyze the model with the least skewed residuals.

go-past time) at the critical word (*a bus* or *a car*) there should be a significant interaction such that loanwords are read faster than non-loanwords in the bilingual readers but not in the monolingual readers. (If loanwords are read faster in both groups, that would mean that the difference is not due to their loanword status—since monolingual English speakers don't have any mental representation of the Cantonese loanword—but rather due to some other factor that is confounded with loanword status.) These two possible predicted patterns are illustrated in the figures below:

Commented [PS[15]: Here I describe the detailed predictions





References

Baayen, H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R.*

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

Bice, K., & Kroll, J. (2015). Native language change during early stages of second language learning. *NeuroReport, 26*, 966-971.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1283-1286.

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.

Van Assche, E., Duyck, W., Hartsuiker, R., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science, 20*, 923-927.