second language research

Research Note

Second Language Research
2021, Vol. 37(4) 681–695

DOI: 10.1177/0267658319877666
journals.sagepub.com/home/slr
SAGE

Language proficiency is only part of the story: Lexical access in heritage and non-heritage bilinguals

Seth Wiener
Carnegie Mellon University, USA

Natasha Tokowicz
University of Pittsburgh, USA

Abstract

This study examined how language proficiency and age of acquisition affect a bilingual language user's reliance on the dominant language during lexical access. Two bilingual groups performed a translation recognition task: Mandarin–English classroom bilinguals who acquired their dominant language (Mandarin) from birth and their non-dominant language (English) post-puberty through formal classroom instruction, and Mandarin–English heritage bilinguals who acquired their non-dominant language (Mandarin) at home from birth but became more dominant in another language (English) through society and peers. Participants decided whether word pairs were correct translations as quickly and accurately as possible. Critical trials involved correct translations (e.g. 房东 – landlord) and incorrect translations that were related to the correct translation in meaning (e.g. 房东 – rent) or form (e.g. 房东 – lantern). When identifying correct translations, lower proficiency heritage bilinguals were slower and less accurate than higher proficiency classroom bilinguals. Yet, when rejecting incorrect translations, heritage bilinguals demonstrated a greater magnitude of semantically-related interference than classroom bilinguals. Heritage bilinguals additionally demonstrated small but measurable amounts of form-related interference whereas the classroom bilinguals did not. Heritage bilinguals thus showed unique patterns of lexical access distinct from bilinguals who acquired their non-dominant language at a later age in a classroom setting.

Keywords

bilingualism, heritage learners, lexical access, Mandarin Chinese, translation recognition

Corresponding author:
Seth Wiener, Carnegie Mellon University, 160 Baker Hall, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
Email: sethw1@cmu.edu

I Introduction

The first-learned language must play a role during the acquisition of a second-learned language. This much is uncontroversial. Interaction between first-learned and second-learned languages is universally instantiated in models of bilingual word recognition and memory (e.g. Dijkstra and Van Heuven, 2002; Grosjean, 1988; Shook and Marian, 2013). As an example, the Revised Hierarchical Model (Kroll and Stewart, 1994) states that beginning bilinguals initially access the meaning of words in the second-learned language through the first-learned language's translation equivalent. Yet, bilingual lexical development may be affected by more than just the order in which languages are learned. At what age a language is acquired (e.g. Hernandez et al., 2005) as well as language proficiency or dominance, which can shift over time (e.g. Polinsky and Kagan, 2007), could both theoretically impact bilingual lexical processing. In the present study, we use the translation recognition task to examine how language proficiency and age of acquisition affect lexical access in two groups of cross-script Mandarin Chinese (hereafter 'Mandarin') and English bilinguals. Novel to our approach, we test a population of heritage language bilinguals who were exposed to a minority language from birth – Mandarin – but became more proficient in a societally dominant language: English. We compare these heritage bilinguals to a population of Mandarin–English bilinguals who also first acquired Mandarin from birth and later acquired English through formal classroom learning. Together, these two groups allow us to examine whether heritage and non-heritage bilinguals go through the same lexical access process despite their dissimilar language learning experiences, proficiency levels, and ages of acquisition.

II Translation recognition task: Insight into the bilingual lexicon

In the translation recognition task (De Groot, 1992; De Groot and Comijs, 1995), a word in one language is briefly displayed on a computer screen, followed by a word in another language. Participants are instructed to indicate via button press as quickly and accurately as possible whether the two words are translation equivalents. This task's linking hypothesis is that the response time it takes to reject an incorrect translation pair, and the resulting accuracy, reveal the architecture of the bilingual lexicon: incorrect translation pairs provide insight into how the two languages interact during lexical access.

As an example, Talamas et al. (1999) tested Spanish–English bilinguals on correct translation pairs (e.g. man – *hombre*) and three different types of incorrect translations: pairs that were related by lexical form (e.g. man – *hambre* ('hunger')), pairs related by meaning (e.g. man – *mujer* ('woman')), and unrelated pairs serving as controls. Talamas et al. demonstrated that participants were slower and less accurate at rejecting the incorrect semantic and form-related pairs than unrelated controls. Additionally, the authors found an asymmetry in interference patterns: lower proficiency bilinguals showed high form-related interference and low semantically-related interference whereas higher proficiency bilinguals showed low form-related interference and high semantically-related interference. Thus, for form-related interference, lower proficiency bilinguals demonstrated asymmetrically more interference than higher proficiency bilinguals. In contrast, for semantically-related interference, higher proficiency bilinguals demonstrated asymmetrically more interference than lower proficiency bilinguals.

Yet, this asymmetry is not always observed in translation recognition studies. For example, Sunderman and Kroll (2006) expanded Talamas et al.'s findings by testing native English speaking adults still acquiring Spanish as their second-learned language through classroom instruction (see also Linck et al., 2009). Whereas Sunderman and Kroll replicated the asymmetry for form-related pairs, the authors did not replicate the asymmetry for semantically-related pairs: lower proficiency bilinguals and higher proficiency bilinguals had similar levels of semantic interference. Sunderman and Priya (2012) tested cross-script Hindi-English bilinguals and failed to replicate the asymmetry for form-related pairs: higher proficiency bilinguals exhibited high form-related interference. Guo et al. (2012) and Ma et al. (2017) both tested cross-script Mandarin–English bilinguals and failed to observe an asymmetry: participants exhibited similar high levels of form-related and semantically-related interference. These mixed results thus challenge Talamas et al.'s (1999) claim that as proficiency increases, lexical links are relied on less and conceptual links are strengthened thereby allowing learners to directly access concepts (i.e. conceptually mediate).

These translation recognition findings are situated within two broader theoretical questions related to bilingual lexical processing: (1) is conceptual mediation possible at differing levels of bilingual proficiency (e.g. Altarriba and Mathis, 1997; Ferré et al., 2006; Kroll and Tokowicz, 2001); (2) does reliance on lexical links change as a function of proficiency in a second-learned/non-dominant language (e.g. Thierry and Wu, 2007; Wu et al., 2013). In the present study, we extend research into language users' reliance on the translation equivalent by examining not only how language proficiency affects lexical access in bilinguals but also how the age of acquisition may impact bilingual lexical processing.

III Bilingualism through different contexts: Heritage and classroom language learning

Heritage language learners represent a growing population of language learners, in large part due to increased immigration and globalization (for reviews, see Montrul, 2010, 2015). These learners are typically characterized as children of immigrants born in the host country and therefore exposed to a minority language, often at a very young age, through different linguistic environments. Yet, many heritage bilinguals undergo a switch in language dominance at some point in their lives and reach adulthood with a lower proficiency in the minority language and a higher proficiency in the societally dominant language (Polinsky and Kagan, 2007). This unique language history makes heritage learners an important test group to compare against non-heritage learners who acquire their second-learned language later in life through structured classroom learning and typically do not undergo a switch in language dominance (e.g. Au et al., 2002; Brinton et al., 2017; Kan and Schmid, 2019).

In this study, we examine Mandarin–English heritage bilinguals who first learned Mandarin at home but failed to develop their proficiency in Mandarin at the same rate as their proficiency in English, the latter of which was developed as a child through society and peers. We also examine non-heritage Mandarin–English classroom bilinguals who became more proficient in their first-learned, societally dominant language (Mandarin) and less proficient in their second-learned language (English), which they acquired post-puberty through structured classroom instruction. Together, these heritage and classroom bilinguals allow us to disentangle whether reliance on the more dominant language during lexical access is solely driven by language proficiency, or if an early age of acquisition within a naturalistic language environment can also affect access to the translation equivalent.

If proficiency primarily drives lexical access (e.g. Talamas et al., 1999), then the more proficient classroom bilinguals should demonstrate high semantically-related interference and low form-related interference while the less proficient heritage bilinguals demonstrate the opposite asymmetry: low semantically-related interference and high form-related interference. If, however, learning a language at a young age results in conceptually-mediated lexical access, then there should be no difference between the two bilingual groups: heritage bilinguals may perform in a manner similar to non-heritage classroom bilinguals despite the heritage bilinguals achieving a much lower proficiency in their non-dominant language. If heritage bilinguals' unique early language exposure to both languages results in strong conceptual links, heritage bilinguals may show even greater semantically-related interference than the classroom bilinguals. Under this account, heritage bilinguals may be able to directly access conceptual information from words in a manner distinct from far more proficient classroom bilinguals.

IV Experiment

1 Participants

Twenty self-identified native Mandarin speakers from mainland China took part as the Mandarin–English classroom bilingual group (hereafter 'classroom bilinguals'). All participants reported first learning and speaking only Mandarin at home. All participants had completed up to high school education in China, and began acquiring English as a second language in a classroom setting in China as early teenagers. At the time of testing, all classroom bilinguals had been living in the USA for a minimum of one year.

Twenty self-identified heritage Mandarin speakers from the USA took part as the Mandarin–English heritage group (hereafter 'heritage bilinguals'). All participants were born in the USA, children of one or more parents born in China, exposed to some variety of spoken Mandarin at home from birth as their first-learned language, and exposed to English as their second-learned language beginning in school (e.g. pre-school/kindergarten). All heritage bilinguals reported regularly hearing and speaking Mandarin with some family members and visiting China previously. At the time of testing, all heritage bilinguals were enrolled in an advanced Mandarin as a foreign language course primarily focused on reading and writing.

All 40 participants across both groups were undergraduate or graduate students at an American university, had normal or corrected-to-normal vision, no history of language disorders, and were familiar with simplified Chinese characters. Participants in both groups completed the Language Experience and Proficiency Questionnaire (LEAP-Q;

Table 1. Objective and self-assessed participant data. Numbers represent means and standard deviations (* indicates a significant difference between the groups using two-sample t-tests, $p < .05$).

	Classroom bilinguals	Heritage bilinguals
First-learned language	Mandarin	Mandarin
Second-learned language	English	English
Dominant language	Mandarin	English
Non-dominant language	English	Mandarin
Sex	15 F; 5 M	12 F; 8 M
Age (years)	20.6 (1.5)	19.7 (0.8)
O-span (percentage correctly recalled)	87.5% (7.1)	91.8% (8.3)
Stroop (degree of interference)	.18 (.08)	.13 (.07)
Self-rated dominant language proficiency (0–10):		
Speaking	9.7 (0.5)	9.8 (0.4)
Comprehension	9.5 (0.5)	9.8 (0.4)
Reading	9.8 (0.4)	9.7 (0.5)
Years of formal non-dominant language education*	10.6 (4.2)	4.8 (2.5)
Self-rated non-dominant language proficiency (0–10):		
Speaking*	7.4 (1.4)	5.9 (2.4)
Comprehension*	8.3 (1.1)	6.6 (2.6)
Reading*	8.2 (1.1)	5.6 (2.5)

Marian et al., 2007) in English as a measure of language proficiency (see Table 1). Participants did not differ in cognitive measures (see below), age, or self-rated dominant language proficiency ratings (ps $> .05$) but did differ in mean years of formal non-dominant language education and self-rated non-dominant language proficiency ratings (ps $< .05$). All participants were paid for their participation.

2 Cognitive measures

Bilingual learners vary in their cognitive abilities, which may relate to their performance on translation tasks; this has been indexed by a variety of individual difference measures including working memory measures and those intended to examine the ability to focus on the task at hand (e.g. Linck et al., 2014; Tokowicz, 2014). Therefore, two additional cognitive measures commonly used within the translation recognition literature were used to control participant variability both within and between each group. After completing the LEAP-Q, participants performed the Operation-span (O-span) task (Engle et al., 1992; Turner and Engle, 1989) and color-word Stroop task (Stroop, 1935). These tasks were presented in English and counterbalanced in testing order.

The O-span task served as a measure of working memory. Participants were shown a fixation for 1,000 ms followed by an operation along with a solution (e.g. $2 * 5 – 2 = 3$) for 2,500 ms. Participants had to indicate via button press whether the solution was correct or incorrect. Following the response or a 1,250 ms timeout, a one- or two-syllable high frequency English word appeared for 1,250 ms (e.g. ‘money’). After each set of operations and words, participants were instructed to recall as many words as they could using the keyboard. Word sets were presented in increasing size from two to six (three sets each) for a total of 60 words. Two practice sets were included. All participants answered over 50% of the operations correctly. Participants' scores represent the mean percentage of total number of words correctly recalled (Table 1). Scoring did not consider word order or operation accuracy. The two groups did not differ in mean number of words correctly recalled $t(38) = -1.48$, $p = .15$.

The color-word Stroop task served as a measure of the ability to ignore task-irrelevant information. Participants were presented with a letter string displayed in color ink and instructed to read aloud the ink color of each stimulus. On each trial the ink color and text were congruent (e.g. the word 'red' presented in red ink), incongruent (e.g. the word 'red' presented in green ink), or neutral (e.g. a string of 'XXXXX' presented in red ink). Items from each condition were presented 20 times in random order for a total of 60 trials. Each trial began with a centered fixation for 1,000 ms, followed by the presentation of the stimulus. The trial advanced after detection of a vocal response. Six practice trials were included (with feedback). Less than 1% of the data were removed for response times (RT) under 300 ms, over 3,000 ms, and incorrect trials. All remaining data were used to calculate an interference score: (incongruent RT – mean[congruent RT + neutral RT]) / mean[RT]. Higher scores indicate greater interference in incongruent trials (Table 1). The two groups did not differ in mean interference, $t(38) = 1.85$, $p = .08$.

3 Materials

A total of 480 translation word pairs were constructed for each participant group. All English and Mandarin words used in the study were nouns. One hundred sixty of the pairs were correct translation pairs (e.g. landlord – 房东), whereas the remaining 320 were incorrect translation pairs. The incorrect pairs were semantically related to the correct translation (e.g. landlord – 租金 ('rent')), related in lexical form to the correct translation (e.g. landlord – 方式 ('style')), or unrelated to the correct translation, i.e. two control condition matched to each incorrect pair (e.g. landlord – 题目 ('topic'); landlord – 负担 ('burden')). The unrelated control items shared neither form nor phonology with the correct translation. Whereas the presentation order of the two words differed between the groups based on the participants' dominant language, the correct pairs, incorrect semantic pairs, and unrelated pairs all used the same items for the two groups. However, to ensure that the nature of the lexical interference was maintained across the two groups, the form-related pairs differed as a function of the participants' dominant language. For example, the classroom bilinguals saw trials in which the translation of the English word (e.g. 'landlord' translation 房东) was related in form to the Mandarin target (e.g. 方式 meaning 'style') through shared initial phonological and form overlap with the correct translation's first character. In contrast, the heritage bilinguals saw trials in which the translation of the Mandarin word (e.g. 房东 translation 'landlord') was related in form to the English target (e.g. 'lantern') through shared initial phonological and form overlap with the correct translation's first few phonemes.

Half of the 480 items were taken from Guo et al. (2012) whereas 240 items in this study differ from those of Guo et al. in two key ways. First, unlike Guo et al.'s stimuli, which contained both disyllabic and monosyllabic items, our stimuli exclusively

Table 2. Means (and SD) for related/unrelated distractors.

	Semantic		Form-related	
	Frequency[a]	Length[b]	Frequency[a]	Length[b]
Mandarin stimuli:				
Related	2.99 (.57)	15.8 (3.2)	3.13 (.65)	15.6 (3.1)
Unrelated	2.96 (.71)	15.1 (2.9)	2.95 (.67)	16.2 (3.3)
English stimuli:				
Related	3.12 (.58)	6.1 (2.1)	3.22 (.81)	5.6 (1.8)
Unrelated	3.01 (.66)	6.2 (2.2)	3.10 (.73)	6.1 (2.3)

Notes. [a]Log10 of the word frequency taken from SUBTLEX-CH (Cai and Brysbaert, 2010) or SUBTLEX-US (Brysbaert and New, 2009). Means did not differ between related/unrelated stimuli or between semantic/form-related conditions (*p*s $>$ .1). [b]Length measures number of strokes (Mandarin) or letters (English) in word. Means did not differ between related/unrelated stimuli or between semantic/form-related conditions (*p*s $>$ .1).

contained disyllabic Mandarin words. Because the majority of spoken Mandarin words by type are disyllabic (Duanmu, 2007; Packard, 2000), this served as a more ecologically valid test of Mandarin lexical access in line with theoretical proposals of the Mandarin lexicon (e.g. Packard, 1999; Zhou and Marslen-Wilson, 1994). We also note that testing monosyllabic morphemes (e.g. 狼 (wolf) and 浪 (wave)) may have exaggerated Guo et al.'s form-related effect given the lack of a disambiguating second morpheme and the need for participants to attend exclusively to phonetic and semantic cues or 'radicals' in the individual characters (for discussions, see DeFrancis, 1986; Perfetti and Tan, 1998).

Second, we replaced Guo et al.'s pairs that used identical initial characters in form-related distractors (e.g. 新郎 (bride) and 新聞 (news), both of which contain the initial 新 morpheme meaning 'new'). This ensured that all form-related distractors were disyllabic compounds in which the first morpheme shared partial overlap but both morphemes were unique.

To control for translation ambiguity, critical items were selected from Wen and van Heuven (2017) and Tseng et al. (2014), when possible. However, due to the unique nature of the stimuli, it was necessary to go beyond the stimuli provided in those norms. Table 2 summarizes the lexical characteristics of the items. Neither word frequency nor word length differed between conditions or related/unrelated items (*p*s $>$ .1). Participants responded to 160 correct and 160 incorrect trials (40 semantic, 40 form, 80 unrelated) for a total of 320 trials across two blocks. In other words, participants all saw the same correct (i.e. 'yes') trials, but the four types of incorrect (i.e. 'no') trials were counterbalanced across lists and presented only once as either a distractor or matched control; for example stimuli, see Appendix in Supplemental material.

4 Procedure

Participants were randomly assigned to one of five different lists, which counterbalanced the presentation of incorrect trials. Participants were shown a fixation cross at the center of the screen for 2,000 ms, followed by a 200 ms ISI, and presentation of the word in the non-dominant language for 500 ms. A 250 ms ISI followed and then the word in the dominant language was displayed for 500 ms. In other words, the heritage bilinguals saw Mandarin–English pairs while the classroom bilinguals saw English–Mandarin pairs. Trials advanced to the fixation cross after each button press. Placement of 'yes' and 'no' buttons were counterbalanced across participants. The O-span, Stroop, and translation recognition tasks were all presented using E-prime 2.0 (Psychology Software Tools, Inc.). Response times were recorded using the Chronos response and stimulus device (Babjack et al., 2015). Participants performed five practice trials (with feedback) prior to the start of the experiment. The entire testing procedure lasted approximately one hour.

V Results

Ninety-five percent confidence intervals for correct 'yes' translation pairs were first calculated for each bilingual group. In all response time (RT) analyses reported here, only correct trials were used. RTs faster than 300 ms or slower than 3,000 ms were removed (less than 4% of all data). Classroom bilinguals overall performed the task faster [536, 570] (ms) and more accurately [.93, .97] than heritage bilinguals [900, 996] (ms) and [.83, .88], respectively. To test whether accuracy and correct RT were statistically different between groups, logistic (accuracy) and linear (correct RT) mixed-effects models were built using the *lme4* package (Bates et al., 2015) and the *lmerTest* package (Kuznetsova et al., 2017) in R version 3.3 (R Core Team, 2017). R code, variable coding, and full model output are available in the Appendix (see Supplemental material). Statistical analyses confirmed that classroom bilinguals were faster (β = –222.83, SE = 35.59, t = –6.26, $p < .001$) and more accurate (β = 1.13, SE = 0.23, z = 4.85, $p < .001$) than heritage bilinguals. Cognitive measures did not affect accuracy or RT in either group (ps $> .05$).

Table 3 reports 95% confidence intervals for each type of related and unrelated 'no' distractor trial. Summary statistics in Table 3 indicate that both bilingual groups were overall slower and less accurate when responding to semantic distractors than form-related distractors. To test whether accuracy and correct RT of 'no' trials differed by group and distractor type, linear and logistic regression models were built. Whereas a main model was built to analyse the semantic and its unrelated controls trials, separate models were built to analyse the form-related distractor items because these items differed as functions of each group's dominant language (for additional model details, see Appendix in Supplemental material).

The semantic trials were responded to less accurately (β = –1.00, SE = 0.09, z = –10.66, $p < .001$) and slower (β = 89.43, SE = 13.98, t = 6.39, $p < .001$) than unrelated controls. This pattern was consistent across both bilingual groups (ps $< .01$) and not affected by cognitive measures (ps $> .05$). Overall, classroom bilinguals exhibited less semantic accuracy (β = 0.52, SE = 0.13, z = 4.01, $p < .001$) and RT (β = –285.30, SE = 46.77, t = –6.10, $p < .001$) interference than heritage bilinguals.

The form-related trials were responded to with similar RT as compared to the unrelated controls. This pattern was observed for both groups (classroom bilinguals: β = –10.60, SE = 6.48, t = –1.64, p = .10; heritage bilinguals: β = –17.48, SE = 19.78,

Table 3. Ninety-five percent confidence intervals for mean accuracy and mean correct response time (ms).

	Semantic		Form-related	
	Accuracy	Response time	Accuracy	Response time
Classroom bilinguals:				
Related	[.77, .83]	[701, 771]	[.92, .96]	[613, 657]
Unrelated	[.95, .97]	[621, 661]	[.95, .97]	[595, 629]
Heritage bilinguals:				
Related	[.60, .68]	[1,236, 1,390]	[.81, .87]	[1,100, 1,220]
Unrelated	[.90, .92]	[1,053, 1,165]	[.89, .91]	[1,082, 1,198]

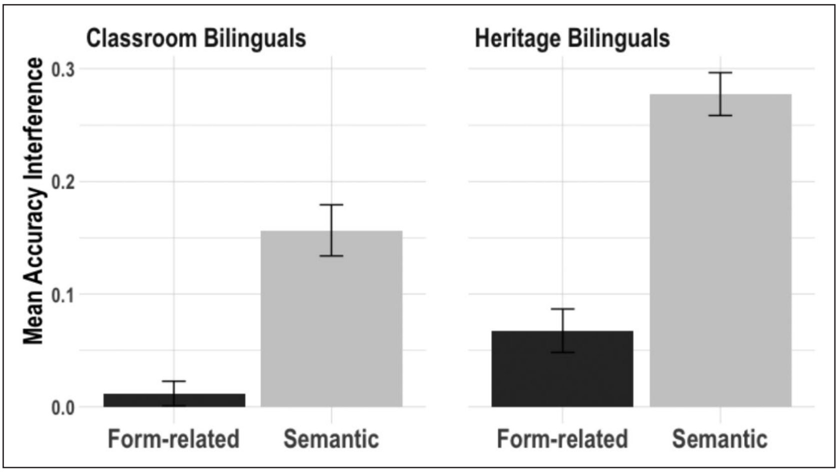

Figure 1. Mean accuracy interference by group and distractor type based on the raw data.

t = –0.88, p = .37). A form-related accuracy interference difference, however, was found; classroom bilinguals showed no accuracy interference (β = –0.20, SE = 0.18, z = –1.11, p = .26) whereas heritage bilinguals responded to form-related distractors less accurately than unrelated controls (β = –0.38, SE = 0.15, z = –2.48, p = .01).

Figure 1 plots the mean accuracy interference for each group based on the raw data from 'no' distractor trials; error bars represent 1 standard error. Figure 2 plots the mean response time interference. Figures 1 and 2 visualize the three interference results from this study: (1) Both groups exhibited an asymmetric pattern of high semantically-related interference and low form-related interference; (2) Overall, heritage bilinguals demonstrated a greater magnitude of semantically-related interference than classroom bilinguals; (3) For heritage bilinguals, form-related distractors had a significant effect on accuracy but not on RT. No such pattern was observed for classroom bilinguals.

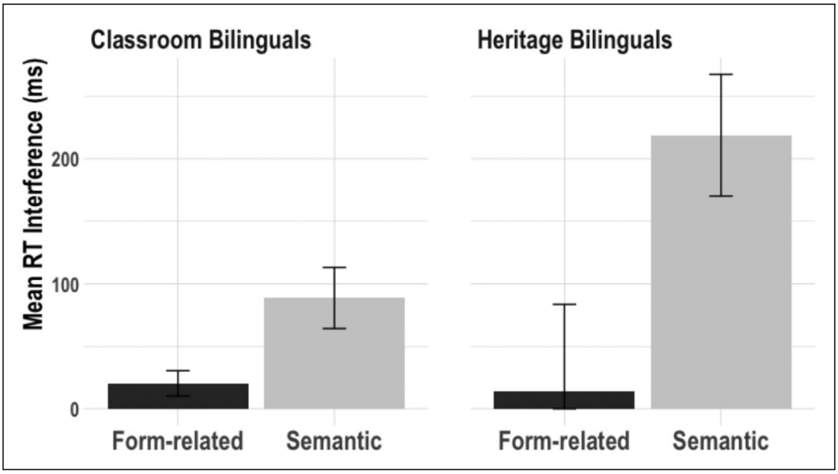

Figure 2. Mean response time (RT) interference by group and distractor type based on the raw data.

VI Discussion

This study made use of the translation recognition task to examine lexical access in two groups of cross script English–Mandarin bilinguals. We examined classroom bilinguals who acquired their second-learned language (English) post-puberty in a structured classroom setting after acquiring their first-learned and dominant language (Mandarin) at home and through society. We additionally examined heritage bilinguals who acquired their first-learned language (Mandarin) at home from birth but ultimately became more proficient in the societally dominant language (English), which they acquired second through society and peers at a young age.

Analyses of the correct translation 'yes' trials revealed a clear behavioral difference between the two groups: classroom bilinguals were faster and more accurate than heritage bilinguals. These results suggest that proficiency in the non-dominant language modulated the processing of correct translation pairs in line with previous translation recognition studies (e.g. Sunderman and Kroll, 2006; Talamas et al., 1999). We note, however, that this proficiency difference may better reflect the immersed classroom bilinguals' superior English abilities and potential cognitive advantage of performing the task in their immersed non-dominant language (see Linck et al., 2009) rather than the heritage bilinguals' less proficient Mandarin abilities. Whereas Guo et al. (2012) did not report correct 'yes' translation accuracy or RTs, our classroom bilinguals appeared to respond faster (mean = 620 ms), more accurately (mean = .95), and have higher self-rated non-dominant language levels than those bilingual participants tested in Guo et al. (based on Guo et al.'s Figures 1 and 2). Moreover, our heritage bilinguals' accuracy (mean = .85) was comparable to that of previously tested (non-heritage) English–Spanish bilinguals deemed 'more proficient' (compare mean = .87 in Sunderman and Kroll, 2006) and those immersed in their second language in a study abroad program (compare mean = .88 in

Linck et al., 2009). These similar accuracies are especially notable given that the Chinese writing system is far more opaque than that of Spanish (DeFrancis, 1986).

Yet, while proficiency in a non-dominant language may partly account for the correct translation results, proficiency alone cannot account for the incorrect translation results. Neither Mandarin–English bilingual group demonstrated a significant level of form-related RT interference. This finding is in conflict with Guo et al.'s (2012) results, which reported roughly similar levels of semantic and lexical interference among high proficiency Mandarin–English bilinguals. This difference between Guo et al.'s results and ours may reflect the nature of the stimuli. In particular, Guo et al. tested words with monosyllabic morphemes (e.g. 妈 'mother' and 马 'horse'), which could have exaggerated the form-related interference. As we outlined in the methods section, the lack of a disambiguating second morpheme and the need for participants to attend exclusively to phonetic and semantic radicals in the individual characters may have delayed responses. Also (although this seems highly unlikely), it bears mentioning that the two studies used different methodologies; specifically, Guo et al. recorded electroencephalographic (EEG) responses during the translation recognition task to measure event-related potentials whereas we did not. It is difficult to envision, however, what about the application of the EEG cap or administration of the procedure could have done to change the nature of the findings, particularly because our procedures were very similar in many other respects.

Unique to the heritage bilinguals, a small but significant (mean = .06) effect of form-related accuracy interference was observed. No such effect was found for the classroom bilinguals (mean = .01). It is unclear whether this accuracy interference reflects differences in non-dominant language proficiency, the context/age of the language learning experience, or simply the challenges associated with acquiring Mandarin literacy skills as dominant English speakers/readers (e.g. Ke and Koda, 2017). We are in the process of collecting translation recognition data from lower proficiency, non-heritage Mandarin learners to clarify this latter point.

Lastly, we observed high degrees of semantic accuracy and RT interference in both bilingual groups (Figure 1). We found that heritage bilinguals, however, demonstrated a larger magnitude of semantic accuracy and RT interference than classroom bilinguals.

Taken together, our results corroborate previous findings that showed asymmetric interference with less form-related interference and greater semantically-related interference for high proficiency bilinguals (e.g. Linck et al., 2006; Talamas et al., 1999). Yet, our results go beyond these previous translation recognition studies by demonstrating that language proficiency is not the only mitigating factor in bilingual lexical access. Heritage bilinguals, despite being far less proficient in their non-dominant language than classroom bilinguals, demonstrated greater semantic accuracy and RT interference than classroom bilinguals. This finding is at odds with predictions from the Revised Hierarchical Model (Kroll and Stewart, 1994), which states that greater proficiency results in a greater reliance on semantics. We speculate that because heritage bilinguals were exposed to both languages at early ages through naturalistic learning contexts, conceptual links were strengthened over a longer period of time without the need for intervening lexical form information. In contrast, classroom bilinguals' conceptual links were relatively weaker given that these bilinguals were exposed to their second-learned language later in life after their first language was already in place.

Moreover, nearly all of our heritage participants self-reported being relatively proficient speakers (and far more proficient than the typical non-heritage classroom second language Mandarin learner) but being unable to read and write Chinese characters until they started university-level Mandarin as a foreign language courses (as is common for most adult heritage Mandarin speakers; for discussions, see Koda et al., 2008; Xiao, 2006). In contrast, all of our classroom bilinguals were literate in Mandarin as early teenagers by the time they began acquiring English in formal classrooms in China. To what degree literacy development may have also affected these results remains unclear. Limited evidence suggests that heritage Mandarin learners demonstrate unique Chinese character/radical trajectories dissimilar from those of non-heritage learners (Chen, 2019).

In sum, we found that non-dominant language proficiency is not the only explanatory factor in bilingual lexical access; the age at which a non-dominant language is acquired can also affect the manner of lexical access in bilinguals. Heritage learners who were exposed to their non-dominant language at home from birth differed from classroom learners who acquired their non-dominant language post-puberty in terms of their reliance on semantically-related and form-related information during translation recognition. These results strengthen the claim that early linguistic experience with a minority language – even if merely overhearing the language – can not only affect language users' perception and production of that language later in life (e.g. Au et al., 2002; Chang, 2016; Chang and Yao, 2016; Chang et al., 2011) but also how users access lexical information in that language. Our results underscore the need to consider a wide range of language learners with different ages of acquisition, proficiency levels, and learning contexts when assessing theoretical models.

Acknowledgements

We thank Katherine Wise and Yvonne Yuanpu Ma for helping us run the study and the incredibly insightful and supportive Second Language Research reviewers for helping us see the bigger (and far more important) picture.

Declaration of Conflicting Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Seth Wiener https://orcid.org/0000-0002-7383-3682

Supplemental material

Supplemental material for this article is available online.

References

Altarriba J and Mathis KM (1997) Conceptual and lexical development in second language acquisition. *Journal of Memory and Language* 36: 550–68.

Au TK, Knightly LM, Jun SA and Oh JS (2002) Overhearing a language during childhood. *Psychological Science* 13: 238–43.

Babjack DL, Cernicky B, Sobotka AJ et al. (2015) Reducing audio stimulus presentation latencies across studies, laboratories, and hardware and operating system configurations. *Behavior Research Methods* 47: 649–65.

Bates D, Mächler M, Bolker B and Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.

Brinton DM, Kagan O and Bauckus S (2017) *Heritage language education: A new field emerging*. Abingdon: Routledge.

Brysbaert M and New B (2009) Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41: 977–90.

Cai Q and Brysbaert M (2010) SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS one* 5: e10729.

Chang CB (2016) Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition* 19: 791–809.

Chang CB and Yao Y (2016) Toward an understanding of heritage prosody: Acoustic and perceptual properties of tone produced by heritage, native, and second language speakers of Mandarin. *Heritage Language Journal* 13: 134–60.

Chang CB, Yao Y, Haynes EF and Rhodes R (2011) Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America* 129: 3964–80.

Chen T (2019) Radical awareness development of L2 Chinese learners in the US. In: Yuan F and Li S (eds) *Classroom research on Chinese as a second language*. New York: Routledge, pp. 155–70.

DeFrancis J (1986) *The Chinese language: Fact and fantasy*. Honolulu, HI: University of Hawaii Press.

De Groot AM (1992) Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18: 1001–18.

De Groot AM and Comijs H (1995) Translation recognition and translation production: Comparing a new and an old tool in the study of bilingualism. *Language Learning* 45: 467–509.

Dijkstra T and Van Heuven WJ (2002) The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5: 175–97.

Duanmu S (2007) *The phonology of standard Chinese*. Oxford: Oxford University Press.

Engle RW, Cantor J and Carullo JJ (1992) Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18: 972–92.

Ferré P, Sánchez-Casas R and Guasch M (2006) Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and nonproficient Spanish-Catalan bilinguals. *Language Learning* 56: 571–608.

Grosjean F (1988) Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes* 3: 233–74.

Guo T, Misra M, Tam JW and Kroll JF (2012) On the time course of accessing meaning in a second language: An electrophysiological and behavioral investigation of translation recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 38: 1165–86.

Hernandez A, Li P and MacWhinney B (2005) The emergence of competing modules in bilingualism. *TRENDS in Cognitive Science* 9: 220–25.

Kan RT and Schmid MS (2019) Development of tonal discrimination in young heritage speakers of Cantonese. *Journal of Phonetics* 73: 40–54.

Ke S and Koda K (2017) Is vocabulary knowledge sufficient for word-meaning inference? An investigation of the role of morphological awareness in adult L2 learners of Chinese. *Applied Linguistics* 40: 456–77.

Koda K, Zhang Y and Yang CL (2008) Literacy development in Chinese as a heritage language. In: He AW and Xiao Y (eds) *Chinese as a heritage language: Fostering rooted world citizenry*. Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center, pp. 137–49.

Kroll JF and Stewart E (1994) Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33: 149–74.

Kroll JF and Tokowicz N (2001) The development of conceptual representation for words in a second language. In: Nicol J (ed.) *One mind, two languages: Bilingual language processing*. Malden, MA: Blackwell, pp. 49–71.

Kuznetsova A, Brockhoff PB and Christensen RHB (2017) lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82: 1–26.

Linck JA, Kroll JF and Sunderman GL (2009) Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science* 20: 1507–15.

Linck JA, Osthus P, Koeth JT and Bunting MF (2014) Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin and Review* 21: 861–83.

Ma F, Chen P, Guo T and Kroll J (2017) When late second language learners access the meaning of L2 words: Using ERPs to investigate the role of the L1 translation equivalent. *Journal of Neurolinguistics* 41: 50–69.

Marian V, Blumenfeld HK and Kaushanskaya M (2007) The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50: 940–67.

Montrul S (2010) Current issues in heritage language acquisition. *Annual Review of Applied Linguistics* 30: 3–23.

Montrul S (2015) *The acquisition of heritage languages*. Cambridge: Cambridge University Press.

Packard JL (1999) Lexical access in Chinese speech comprehension and production. *Brain and Language* 68: 89–94.

Packard JL (2000) *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.

Perfetti CA and Tan LH (1998) The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24: 101–18.

Polinsky M and Kagan O (2007) Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass* 1: 368–95.

R Core Team (2017) *R Foundation for Statistical Computing* [online]. Vienna.

Shook A and Marian V (2013) The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition* 16: 304–24.

Stroop JR (1935) Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18: 643–62.

Sunderman GL and Kroll JF (2006) First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition* 28: 387–422.

Sunderman GL and Priya K (2012) Translation recognition in highly proficient Hindi–English bilinguals: The influence of different scripts but connectable phonologies. *Language and Cognitive Processes* 27: 1265–85.

Talamas A, Kroll JF and Dufour R (1999) From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition* 2: 45–58.

Thierry G and Wu YJ (2007) Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences* 104: 12530–35.

Tokowicz N (2014) *Lexical processing and second language acquisition*. New York: Routledge.

Tseng AM, Chang L-Y and Tokowicz N (2014) Translation ambiguity between English and Mandarin Chinese: The roles of proficiency and word characteristics. In: Schwieter JW and Ferreira A (eds) *The development of translation competence: Theories and methodologies from psycholinguistics and cognitive science*. Cambridge: Cambridge Scholars Publishing, pp. 107–65.

Turner ML and Engle RW (1989) Is working memory capacity task dependent? *Journal of Memory and Language* 28: 127–54.

Wen Y and van Heuven WJ (2017) Chinese translation norms for 1,429 English words. *Behavior Research Methods* 49: 1006–19.

Wu YJ, Cristino F, Leek C and Thierry G (2013) Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition* 129: 418–25.

Xiao Y (2006) Heritage learners in the Chinese language classroom: Home background. *Heritage Language Journal* 4: 47–56.

Zhou X and Marslen-Wilson W (1994) Words, morphemes and syllables in the Chinese mental lexicon. *Language and Cognitive Processes* 9: 393–422.