Cognition 169 (2017) 15–24

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

# Do you hear '*feather*' when listening to '*rain*'? Lexical tone activation during unconscious translation: Evidence from Mandarin-English bilinguals

Xin Wang [a,b,*], Juan Wang [c], Jeffrey G. Malins [d]

[a] University of Greenwich, UK
[b] University of Oxford, UK
[c] Jiangsu Normal University, China
[d] Haskins Laboratories, Yale University, USA

ARTICLE INFO

*Article history:*
Received 8 September 2016
Revised 28 July 2017
Accepted 31 July 2017

*Keywords:*
Unconscious translation
Lexical tone
Bilingualism
Spoken word recognition
Visual world paradigm
Bilingual lexical access

ABSTRACT

Although lexical tone is a highly prevalent phonetic cue in human languages, its role in bilingual spoken word recognition is not well understood. The present study investigates whether and how adult bilinguals, who use pitch contours to disambiguate lexical items in one language but not the other, access a tonal L1 when exclusively processing a non-tonal L2. Using the visual world paradigm, we show that Mandarin-English listeners automatically activated Mandarin translation equivalents of English target words such as '*rain*' (Mandarin '*yu3*'), and consequently were distracted by competitors whose segments and tones overlapped with the translations of English target words ('*feather*', also '*yu3*' in Mandarin). Importantly, listeners were not distracted by competitors that overlapped with the translations of target words in all segments but not tone ('*fish*'; Mandarin '*yu2*'), nor were they distracted by competitors that overlapped with the translations of target words in rime and tone ('wheat', Mandarin '*gu3*'). These novel results demonstrate implicit access to L1 lexical representations through automatic/unconscious translation, as a result of cross-language top-down and/or lateral influence, and highlight the critical role of lexical tone activation in bilingual lexical access.



## 1. Introduction

A unique challenge to a bilingual speaker/listener is the mastery of two linguistic systems that may maintain conflicting rules. One under-explored area of potential conflict lies in the linguistic dimension of supra-segmental information (i.e., lexical tones). Unlike English, a non-tonal language, in which pitch movements may be used to differentiate emotions, questions and statements, or stress and focus (Gussenhoven, 2004), in a tonal language such as Mandarin Chinese, systematic variation in pitch contours needs to be integrated with segmental information in order to disambiguate lexical items. Mandarin has four lexical tones corresponding to four distinct pitch contours, which are typically represented with numerals as in the following example: *ma1* 'mother', *ma2* 'hemp', *ma3* 'horse', *ma4* 'scold' (Chao, 1968). Phonetically, Mandarin tones 1–4 can be described as high level [55], high rising [35], low falling rising [214], and high falling [51], respectively, with the lowest pitch level assigned a value of 1 and the highest level assigned a value of 5 in phonetic transcription (Howie, 1976). Because of this difference in the status of supra-segmental information between languages, bilinguals of one tonal language and one non-tonal language (e.g., Mandarin-English) offer a unique window to understand the interplay between linguistic processing and representation at the supra-segmental level versus the segmental level in bilingual individuals.

How native speakers of tonal languages identify words in speech has attracted a tradition of debate on the relative weighting of segmental versus supra-segmental cues in constraining word recognition (Cutler & Chen, 1997; Lee, 2007; Liu & Samuel, 2007; Malins & Joanisse, 2010, 2012; Schirmer, Tang, Penney, Gunter, & Chen, 2005; Sereno & Lee, 2015; Taft & Chen, 1992; Tong, Francis, & Gandour, 2008; Wiener & Turnbull, 2016; Zhao, Guo, Zhou, & Shu, 2011). Although the view still remains somewhat controversial, strong empirical evidence supports the claim that tone plays a critical role in spoken word recognition, perhaps even equal to that of segmental information. In particular, using the visual

* Corresponding author at: Center for Applied Research and Outreach in Language Education, University of Greenwich, SE10 9LS, UK.
*E-mail addresses:* X.Wang@gre.ac.uk, xinwang267@gmail.com (X. Wang).

http://dx.doi.org/10.1016/j.cognition.2017.07.013
0010-0277/

world paradigm, Malins and Joanisse (2010) reported significant competition effects in conditions in which competitors shared *word-initial segments and tones* (e.g., *chuan2* '*ship*') or *segments only* (e.g., *chuang1* '*window*') with target items (e.g., *chuang2* '*bed*'), but not in conditions in which competitors shared *rimes and tones* (e.g., *huang2* '*yellow*') or *tones* only (e.g., *niu2* '*cow*') with targets. Furthermore, results demonstrated a similar time course of resolution of targets from competitors across the two conditions with word-initial overlap, even though the type of information that distinguished targets from competitors was different across the two conditions – namely, word final phonemes versus lexical tones (e.g., the competitor '*chuan2*' versus the competitor '*chuang1*' in relation to the target '*chuang2*'). This finding indicated that segmental and tonal information are accessed concurrently and play a comparable role in recognizing words in a tonal language.

If tones are crucial in Mandarin spoken word recognition, a key question that arises is whether this linguistic knowledge is utilized during bilingual spoken word recognition. Thus, we are interested in the question of how lexical tones are processed and represented in the bilingual brain. To address this question, we have elected to use the visual world paradigm because it provides temporally sensitive measures of lexical activation and competition during the unfolding of auditory input (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). These measures are collected by recording eye-movements when participants are instructed to pick a target matching an auditory stimulus in an array of pictures on a computer screen (Tanenhaus, Magnuson, Dahan, & Chambers, 2000). The experimental manipulation is realized through the presence of a competitor picture in the array, the name of which bears a phonological relationship with the target, as shown in the above Mandarin example. The logic is that if the auditory input activates the lexical representations associated with competitor pictures, participants are expected to spend time fixating on these pictures during the unfolding of the auditory input and consequently show significant delays in their looks to targets compared to trials in which these competitor pictures are absent. This paradigm has been extensively adopted in bilingual work showing that phonological overlap between targets and competitors can also induce cross-language lexical competition, similar to within-language competition, even in an experiment exclusively conducted in one language (Ju & Luce, 2004; Spivey & Marian, 1999; Weber & Cutler, 2004). For instance, Russian-English bilinguals showed a larger proportion of fixations to inter-lingual competitors (e.g., *marku* '*stamp*') that were phonologically similar to targets (e.g., *marker*) than they did to phonologically unrelated distractors even in an English context (Spivey & Marian, 1999). This evidence was taken to support an input-driven language non-selective mechanism of bilingual lexical access: namely, language co-activation (Wang & Hui, submitted for publication).

Another line of research shows that language co-activation can be not only input-driven, reflecting bottom-up processing, but can also be the result of top-down and/or lateral influences cross-linguistically (Giezen, Blumenfeld, Shook, & Marian, 2015; Shook & Marian, 2012; Thierry & Wu, 2007; Wu, Cristino, Leek, & Thierry, 2013; Wu & Thierry, 2010, 2012; Zhang, van Heuven, & Conklin, 2011). In their seminal work, Thierry and Wu (2007) showed, using ERPs, that Mandarin-English bilinguals implicitly accessed the Mandarin translation equivalents of targets when making semantic relatedness judgments on English pairs (e.g., 邮政-邮件 were activated when judging *post-mail*), as indexed by an amplitude reduction of the N400 component only for target pairs whose Mandarin translations shared first characters (e.g., 邮), compared to pairs whose translations did not share any characters in common. Even though this subtle effect appears to be more detectable in ERP measures compared to behavioral data, as indicated by Wu and Thierry (2010), Wu et al. (2013) demonstrated implicit access to translation equivalents even in a visual search non-linguistic task. More specifically, they showed that Mandarin-English bilinguals fixated more frequently on English words (e.g., *reason*) whose Mandarin translations (e.g., *yuan yin*) began with a morpheme which was a homophone of the Mandarin word for *circle* (i.e., *yuan*) or *square* (i.e., *fang*) when searching for a *circle* or *square* in arrays of English words that sometimes contained one of these shapes. This effect appears to be the result of cross-language lexical competition initiated by phonological activation of translations of the English items from the input. For example, *yuan yin* is the translation of *reason*, which phonologically overlaps the Mandarin word for *circle* (i.e., *yuan*), thus resulting in longer fixations compared to the other items in the array whose translations did not overlap in phonological form with the Mandarin word for *circle*. These findings offered evidence of a top-down and lateral cross-language influence in bilingual non-selective lexical access, as the input did not overlap with the non-target language in form.

At the behavioral level, compared to input-driven language co-activation, implicit access to translation equivalents offers compelling evidence of cross-language interaction in a top-down and/or lateral fashion without cross-language input overlap, but currently remains underexplored. Importantly, it provides a perspective to understand top-down mechanisms in bilingual language processing, and draws from a line of investigations regarding the role of top-down mechanisms in language comprehension in the monolingual literature (e.g., Huettig & Altmann, 2005; Yee & Sedivy, 2006). Therefore, our goal in the present study is to take this perspective to investigate the extent to which top-down activation can influence bilingual spoken word recognition, and the extent to which tonal information contributes to this process. If we observe cross-linguistic tonal activation in a non-target language (i.e. Mandarin) via unconscious/automatic translation, while the input/target language (i.e. English) does not utilize tonal information for word recognition, this will be strong evidence for language co-activation driven by a top-down and/or lateral mechanism rather than bottom-up influence via input overlap. In a visual world paradigm experiment conducted entirely in English, we instructed Mandarin-English bilingual participants to pick target pictures (e.g., *rain* '*yu3*') in an array. Experimental conditions differed with respect to competitor items present on the screen, which occurred in one of three different competitor conditions plus a baseline control condition: *Segmental + Tone*, where the competitor was a homophone of the target's Mandarin translation (e.g., *feather* '*yu3*'); *Segmental*, where the competitor shared all segments with the target translation but differed in tone (e.g., *fish* '*yu2*'); and *Rime + Tone*, where the competitor overlapped in rime and tone with the target translation but differed in onset (e.g., *wheat* '*gu3*'). We hypothesized that (1) language co-activation can occur as a result of top-down and/or lateral activation without any input form overlap; (2) lexical tone is a critical cue in accessing non-target Mandarin words during automatic/unconscious translation. Therefore, we predicted that lexical tone activation via implicit access to translation equivalents would lead to greater competition in the *Segmental + Tone* condition compared to the other conditions, which would manifest as a decreased amount of target fixations as well as an increased amount of competitor fixations.

## 2. Methods

### 2.1. Participants

Testing took place at the Psychology Laboratory at Jiangsu Normal University located in Xuzhou, China. We tested 40 Mandarin-English bilinguals, whose descriptive information is presented in Table 1. As shown in the table, participants were

**Table 1**
Descriptive information concerning the group of subjects ($N = 40$) who performed the eye-tracking experiment.

| Measure | Mean | Range |
|---|---|---|
| Age | 20.7 | 18–24 |
| Age of English acquisition | 10.3 | 8–14 |
| Length of English study | 10.1 | 8–14 |
| English proficiency-CET raw scores[a] | 468.2 | 406–543 |
| English proficiency-CET percentile ranking | 33% | 11–70% |
| Self-rated English listening ability[b] | 3.1 | 2–6 |
| Self-rated English speaking ability[b] | 3.1 | 2–4 |
| Self-rated English reading ability[b] | 3.7 | 2–6 |
| Self-rated English writing ability[b] | 3.6 | 2–5 |

[a] Measured using raw scores on the College English Test (CET-4), an advanced English proficiency test for non-English majors in China. The test scores, in the range of 290–710, index the level of English proficiency: higher scores mean a higher level of proficiency in relation to fellow test-takers.
[b] Self-rating on a scale of 1–7, with 7 being the highest.

moderately proficient in English, as illustrated by their self-ratings and College English Test (CET-4) scores, and all participants learned English in either late childhood or early adolescence. Sample size was decided based on power analysis using the program G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007; a two-sided test for proportions with an odds ratio of 8, a proportion of discordant pairs of 0.3, and a power of 0.85 translates to a sample size of 40). The dataset was initially analyzed after 20 participants' data were collected; after this, the dataset was not analyzed again until all 40 participants had been tested.

### 2.2. Stimuli

Auditory stimuli were all monosyllabic English words, consisting of 35 easily imageable nouns in total, including 7 targets and 28 competitors (See Supplemental Materials for a full list of the stimuli). Altogether, seven sets of items were designed with respect to what will henceforth be termed "critical targets" in English (e.g., *rain*). For each critical target, three competitor conditions were defined based on the relationship between the name of a competitor picture in Mandarin and the Mandarin translation equivalent of the critical target (e.g., *yu3* for the critical target *rain*). In the Segmental + Tone condition, the name of the competitor picture overlapped the Mandarin translation equivalent in all phonemes and tone; in the Segmental condition, the competitor overlapped the translation equivalent in all phonemes but not tone; in the Rime + Tone condition, the competitor overlapped the translation equivalent in vowels and tone but not the onset consonant. For reference, a sample stimulus set is provided in Table 2. All the target words, competitors, and distractors – 91 items in total – were semantically distinctive from each other within languages and across languages; this was necessary, as items were depicted pictorially. In particular, target pictures did not overlap with competitor pictures in visual features, and English target words were semantically unrelated to English words in the competitor conditions. In addition, targets and competitors in English were phonologically/orthographically distinctive without any overlap. Thus, it is only the Mandarin phonological relationship that distinguished competitors from unrelated distractors, which did not bear any relationship with the target in either language.

**Table 2**
Sample stimuli in each of the experimental conditions.

| Experimental condition[a] | Target (English/ Mandarin) | Competitor (English/ Mandarin) |
|---|---|---|
| Segmental + Tone | Rain/*yu3* | Feather/*yu3* |
| Segmental | Rain/*yu3* | Fish/*yu2* |
| Rime + Tone | Rain/*yu3* | Wheat/*gu3* |

[a] Experimental conditions were defined based on the phonological relationship between Mandarin translations of targets and competitors. Importantly, there was no phonological relationship between the English names of targets and the names of the items concurrently presented within the picture arrays.

The mean duration of the auditory stimuli was 647 ms (SD 75 ms) for the seven critical targets, which were the only trials analyzed as outlined below. However, for the sake of counterbalancing, mean duration was matched between targets and competitors and across the three competitor conditions. In addition, for critical targets versus competitors, and across the three competitor conditions, items were balanced for word frequency in both Mandarin and English, as measured using Cai and Brysbaert (2010) as well as the CELEX database (Baayen, Piepenbrock, & van Rijn, 1995).

Each of the 35 English words was digitally recorded as produced by a female adult native speaker of British English, via the open source software Audacity, version 2.0.3 (Audacity Development Team, 2013) at 44.1 kHz. All sound tokens were trimmed and normalized. Line-drawing pictures matching these items were selected from the Google line drawing search database and tested by two Mandarin-English speakers to ensure the selected pictures were prototypical and did not elicit lexical ambiguity in either Mandarin or English. To ensure that listeners were familiar with the words of interest, participants also performed a naming task prior to the eye-tracking study. In this task, participants were presented with each picture and asked to say aloud the monosyllabic Mandarin word and English word that they thought were most appropriate for it. As Mandarin is morphologically productive, a picture can be lexicalized using either a monosyllabic or disyllabic word. If participants named a picture using a disyllabic word, we gave them the intended monosyllabic word. For instance, 'dam' can either be '/da4 ba4/' or '/ba4/' in Mandarin, and so we told participants who named the picture '/da4 ba4/' that the intended name was '/ba4/'. This naming task was adopted to prime listeners' Mandarin knowledge corresponding to the pictures and to ensure the consistency of the intended names of each picture among participants in both languages. However, the drawback of this procedure is that pre-activating Mandarin might have affected the activation level of Mandarin in the subsequent eye-tracking experiment. Although this is a methodological limitation, this should not affect the way we interpret the results, as this pre-activation would have equally applied to all experimental conditions, and we base our conclusions solely on relative differences between experimental conditions.

In addition, to ensure naming and translation consistency for the 91 selected items, including distractors, we first obtained naming data from 20 Mandarin-English bilingual participants who did not participate in the actual experiment. Table 3 shows the breakdown of their naming accuracy across conditions; as shown in the table, there was no significant difference in naming accuracy across experimental conditions or languages, suggesting that the intended names for the pictures were highly consistent across bilingual participants.

### 2.3. Procedure

Participants performed an auditory word-visual picture-matching task during which their eye movements were recorded. On each trial participants were presented with an array of four pictures on a computer screen. Pictures were oriented directly above, below, to the left, or to the right of a fixation point consisting of a black cross (see a sample stimulus array in Fig. 1). On the experimental trials, each array consisted of a target item, a covert translation competitor (in one of three conditions: Segmental + Tone, Segmental, Rime + Tone), and two unrelated distractor items that

**Table 3**
Mean percent accuracy (%) for the naming procedure in both Mandarin and English across the different conditions. Standard Errors are in parentheses.

| | Target | Seg + Tone | Segmental | Rime + Tone | Baseline |
|---|---|---|---|---|---|
| Mandarin | 98.6 (0.02) | 97.1 (0.04) | 97.9 (0.04) | 97.1 (0.08) | 97.9 (0.06) |
| English | 99.3 (0.02) | 98.6 (0.02) | 99.3 (0.02) | 98.6 (0.02) | 98.6 (0.04) |

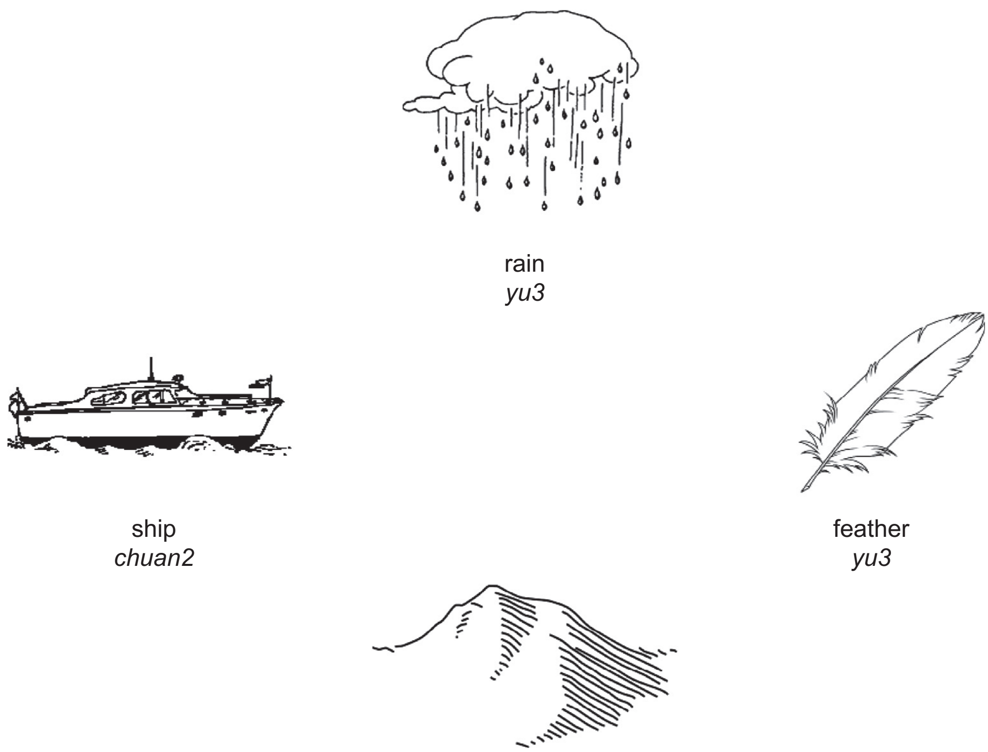


**Fig. 1.** A sample stimulus array. Note that solely pictures were presented to the participants; the names of the pictures in English and in Mandarin are simply provided for reference.

were phonologically unrelated to the critical items and the translations of the critical items. These distractor items were balanced for frequency across competitor conditions and did not overlap phonologically – in either segments or tones – with targets or competitors in both Mandarin and English. Thus, these distractor items, carrying different tones from both targets and competitors, were different across experimental conditions in order to diversify the picture arrays. Baseline trials consisted of target items instead presented with three phonologically unrelated distractor items and thus no competitors. For the purposes of analysis, 'competitor' curves for the baseline condition were calculated by dividing summed fixations to the unrelated distractors by three. The positions of the target and competitor in the display were counterbalanced across trials, as was the relationship between the two pictures in the array (adjacent or opposite), to offset the effect of any influence of target/competitor orientation on eye movement data.

Each trial proceeded as follows: first, a black central fixation cue appeared on the screen, followed by an 800 ms delay; next, an auditory carrier phrase was initiated at the same time as the picture array appeared on the screen. This carrier phrase consisted of the words "I will say..." immediately prior to the presentation of each target word. The "I will say..." token was 2000 ms in duration, was recorded by the same native speaker as the target items, and was bridged with each target word using the open software Audacity version 2.0.3 (Audacity Development Team, 2013), so that listeners were prepared for target items in a natural way. Participants were instructed to look at the fixation cue when it appeared, and to push a button on a keypad that corresponded with the position of the picture matching the word that they heard (↑↓← →). Subsequent trials were initiated either once a response was made, or after 5000 ms had elapsed. Note that participants were not explicitly instructed to look at the picture of the target item.

Participants completed 224 trials. These consisted of 112 experimental trials (i.e., 16 trials for each target item, divided equally across the three experimental conditions, plus the baseline control condition), as well as 112 trials in which the competitor/baseline words were actually the auditory stimuli (i.e., targets). This design, motivated by the study design from Malins and Joanisse (2010), was chosen so there was a 0.50 probability of hearing the name of either the competitors or the baseline distractor to avoid bias towards the target pictures before hearing the auditory stimuli. In addition, as in the Malins and Joanisse (2010) study, the high extent of repetition helped ensure that participants were highly familiar with the items in the experiment, and thus generated strong expectations of upcoming auditory input. Participants were randomly assigned to one of the three lists of pseudo-randomized trials designed to ensure that at least four trials elapsed before the same auditory stimulus or the same array was repeated. Prior to beginning the task, participants also completed a practice block of five trials that did not use any of the 35 items from the actual experiment.

After completing the eye-tracking task, a language history questionnaire was administered to participants. All instructions were provided in English in order to limit any potential expectations that participants' knowledge of Mandarin would be relevant to the experiment.

### 2.4. Acquisition and analysis of eye-tracking data

Eye-tracking data were acquired using a desktop Eye-Link 1000 (SR Research Ltd., Ottawa, Canada) at a sampling rate of 1000 Hz. Participants were presented with arrays of pictures on a 21-in. computer screen located 20 in. away. Prior to data collection, calibration, validation and drift check were performed.

Importantly, we only analyzed trials in which target items were the "critical targets" from which sets of items were derived. Consequently, even though we only analyzed half of the total number of trials that participants performed, this meant that for the trials which we analyzed, all conditions shared the exact same auditory stimuli and only differed with respect to the identity of the competitor picture on the screen. That is, for a given target such as 'rain', trial conditions differed only in competitors ('feather' for the Segmental + Tone condition, 'fish' for the Segmental condition, 'wheat' for the Rime + Tone condition, or three phonologically unrelated distractors for the baseline condition). This procedure allowed us to control for attendant psycholinguistic factors such as duration and frequency of the target items.

Eye-tracking data were analyzed by first resampling fixations to 50 Hz. Following this, fixation proportions were calculated at each time point by summing fixations across the 28 trials in a condition, and dividing the total number of fixations made to a given item type (target, competitor, distractor) by the total number of fixations made to all item types.

Analysis of fixations was restricted to the time window between 200 and 950 ms following the onset of the target stimulus (i.e., 200 ms after the offset of the carrier phrase "I will say..."). The lower limit represents a delay corresponding to the approximate amount of time required to plan and execute an eye movement, whereas the upper limit was defined based on when peak target fixation proportions were reached.

Fixation data were statistically analyzed using growth curve modeling (Mirman, 2014); these procedures were implemented using the *lme4* package (version 1.1-7; Bates, Maechler, Bolker, & Walker, 2015) in R version 3.1.3 (R Core Team, 2014). In these analyses, the baseline condition was used to create a base model, and parameters were estimated for the other three experimental conditions (Segmental + Tone; Segmental; Rime + Tone) relative to baseline. The time course of target fixations was modeled using a third-order orthogonal polynomial with fixed effects of condition and random effects of participants and participants-by-condition. For competitor fixations, a third-order model with the same structure was also used. To test our hypothesis that tonal information is critical for lexical access in Mandarin-English bilinguals, we performed a follow-up analysis in which we created base models for target and competitor fixations in the Segmental + Tone condition and estimated parameters for the Segmental condition relative to this. These models were also third-order with fixed effects of condition and random effects of participants and participants-by-condition.

For all growth curve analyses, statistical significance of individual parameter estimates was assessed using the normal approximation, which treats *t*-values as *z*-values. This is a reasonable approximation when degrees of freedom are relatively large, which they were in this case.

## 3. Results

### 3.1. Behavioral data

Mean accuracy and reaction time for the button press response are presented in Table 4. Reaction times were calculated with respect to the onset of the spoken target word (i.e., after the phrase "I will say"). Only trials with RTs > 250 ms were considered for both accuracy and RT analyses. Furthermore, only correct trials were considered for RT analyses. Both accuracy and RT were analyzed using linear mixed effects models with participants and items treated as random effects. More specifically, a base model was constructed with random intercepts for participants and items, and by-participants and by-item random slopes for the effect of experimental condition. Fixed effects of experimental condition were then added to this model (the accuracy model, which was binomial in nature, was constructed using the program *glmer* in the R package *lme4*, whereas the reaction time model was constructed using the program *lmer* in the R package *lme4*). Improvement in model fit was assessed by change in deviance, which follows a chi-square distribution, with degrees of freedom equal to the number of parameters added to the model. For both the accuracy and reaction time models, there was no significant improvement in model fit after adding fixed effects of experimental condition [accuracy: $\chi^2(3) = 3.039$, $p = 0.386$; reaction time: $\chi^2(3) = 1.093$, $p = 0.779$], suggesting that overt behavioral responses did not significantly differ across experimental conditions. As can be inferred from the table, accuracy was very high for all conditions, suggesting participants did not experience difficulty in performing this task.

### 3.2. Eye-tracking data

#### 3.2.1. Group-wise analysis of target and competitor fixations

Grand average fixation proportions are presented for targets in Fig. 2 and competitors in Fig. 3. Model fits are superimposed over raw data points for each condition. Growth curve parameters for each model are summarized in Table 5. As can be inferred from the figure, the proportion of target fixations in the Segmental + Tone condition showed the greatest difference compared to baseline. Across the entire time window of analysis, target fixation proportions were lower in this condition than baseline, indicative of a competitive effect. Furthermore, the difference between this condition and baseline became more pronounced over time. These observations are respectively reflected in the growth curve parameters; the Segmental + Tone condition showed a significantly smaller intercept than the baseline condition as well as a marginally significantly shallower slope. In addition to this, the Rime + Tone condition showed a marginally significantly shallower slope compared to baseline. The segmental condition did not show significant differences from the baseline model in any model component.

Because we only analyzed trials with critical targets, the only difference between conditions was in the identity of the

**Table 4**
Mean reaction time (ms) and mean percent accuracy (%) for the button press response relative to word onset. Standard Errors are in parentheses.

| Condition | Reaction time (SE) | Percent accuracy (SE) |
|---|---|---|
| Segmental + Tone | 1041 (31) | 95.5 (0.94) |
| Segmental | 1021 (29) | 96.3 (0.73) |
| Rime + Tone | 1052 (32) | 97.1 (0.80) |
| Baseline | 1028 (33) | 97.4 (0.86) |

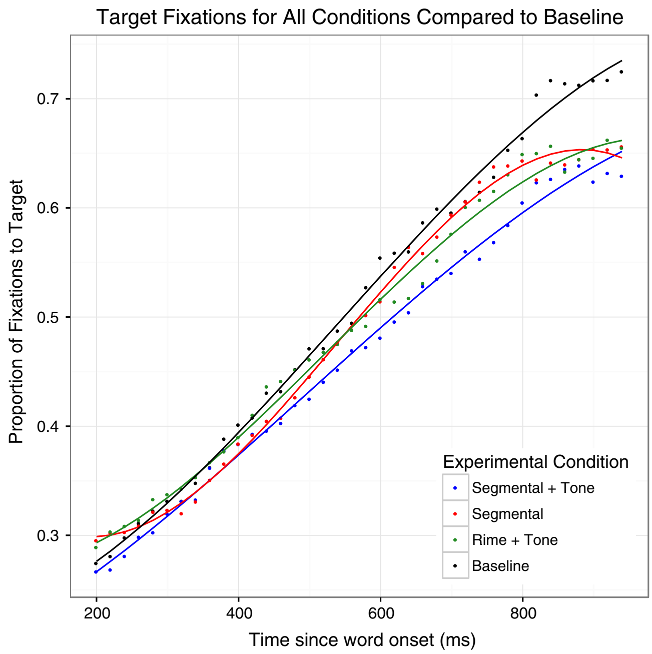


**Fig. 2.** Grand average fixations to targets across participants and items. The points represent mean proportions of target fixations at each time step, whereas the lines represent third order growth curve model fits.

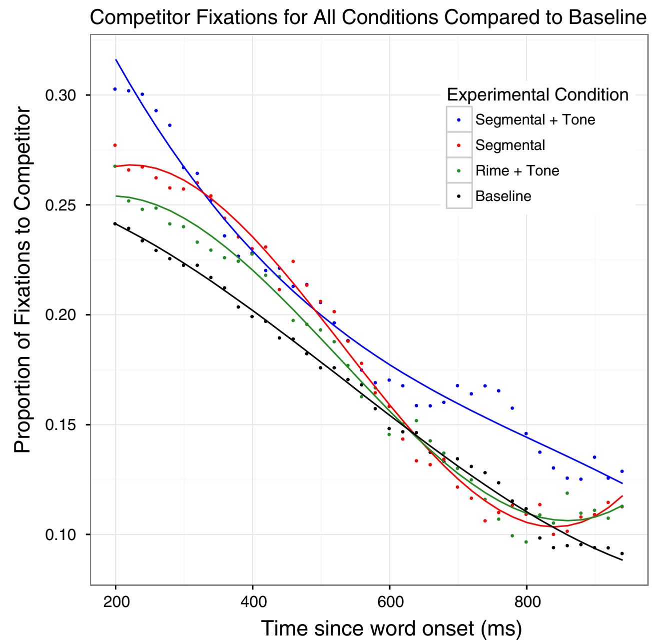


**Fig. 3.** Grand average fixations to competitors across participants and items. The points represent mean proportions of competitor fixations at each time step, whereas the lines represent third order growth curve model fits.

competitor picture on the screen. Therefore, one can infer that target fixations were lower in the Segmental + Tone condition because participants were more inclined to fixate competitor items. As can be seen in Fig. 3, where grand average competitor fixations are displayed along with models fits for each competitor condition relative to baseline, this was the case. The Segmental + Tone condition showed a significantly larger intercept than the baseline condition, whereas the other two conditions were not significantly different from baseline in any model fit component.

### *3.2.2. Direct comparison of the segmental and tonal overlap condition versus segmental overlap alone*

As we were specifically interested in whether tonal information is critical for activation of L1 items in the lexicon in Mandarin-English bilinguals, we used growth curve modeling to directly compare fixations in the Segmental + Tone versus Segmental conditions for both targets and competitors. In these models, the Segmental + Tone condition was used to create a base model, and parameters for the Segmental condition were estimated relative to this base model. The results of this analysis are shown in Figs. 4 and 5, and growth curve parameters are summarized in Table 6.

For both targets and competitors, the Segmental condition showed a significant difference in the cubic component compared to the Segmental + Tone condition; this difference was especially prominent between 500 and 800 ms.

### *3.2.3. Individual difference analyses*

Last, we wished to test whether competitive effects were related to individual differences in age of acquisition and proficiency of L2 English. To do this, we quantified the random effects components for each participants in the target and competitor growth curve models, and correlated these values with two measures: (1) age of acquisition of English and (2) English proficiency as measured using raw CET-4 scores. To limit the number of correlations we performed, we only focused on differences which were significant at the group level. Namely, we assessed individual differences in intercepts between the Segmental + Tone condition and the baseline condition in each of the target and competitor fixation models.

As shown in Fig. 6, there was a significant correlation between L2 proficiency and the size of the difference between the intercept

**Table 5**
Growth curve analysis results for the target and competitor fixation models across all experimental conditions.

| Component | Targets | | | | Competitors | | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | SE | *t* | *p* | Estimate | SE | *t* | *p* |
| Intercept – Segmental + Tone | −0.044 | 0.017 | −2.55 | 0.011 | 0.034 | 0.010 | 3.31 | 0.001 |
| Intercept – Segmental | −0.022 | 0.017 | −1.26 | 0.206 | 0.016 | 0.010 | 1.55 | 0.121 |
| Intercept – Rime + Tone | −0.020 | 0.017 | −1.15 | 0.248 | 0.009 | 0.010 | 0.932 | 0.351 |
| Linear – Segmental + Tone | −0.157 | 0.087 | −1.80 | 0.071 | −0.032 | 0.059 | −0.543 | 0.587 |
| Linear – Segmental | −0.109 | 0.087 | −1.25 | 0.212 | −0.072 | 0.059 | −1.22 | 0.221 |
| Linear – Rime + Tone | −0.147 | 0.087 | −1.69 | 0.091 | −0.029 | 0.059 | −0.483 | 0.629 |
| Quadratic – Segmental + Tone | −0.007 | 0.064 | −0.108 | 0.914 | 0.064 | 0.051 | 1.25 | 0.213 |
| Quadratic – Segmental | −0.034 | 0.064 | −0.535 | 0.593 | 0.036 | 0.051 | 0.705 | 0.481 |
| Quadratic – Rime + Tone | −0.019 | 0.064 | −0.294 | 0.769 | 0.028 | 0.051 | 0.547 | 0.585 |
| Cubic – Segmental + Tone | 0.017 | 0.050 | 0.330 | 0.741 | −0.031 | 0.037 | −0.835 | 0.404 |
| Cubic – Segmental | −0.072 | 0.050 | −1.43 | 0.153 | 0.058 | 0.037 | 1.59 | 0.112 |
| Cubic – Rime + Tone | −0.013 | 0.050 | −0.264 | 0.792 | 0.038 | 0.037 | 1.03 | 0.302 |

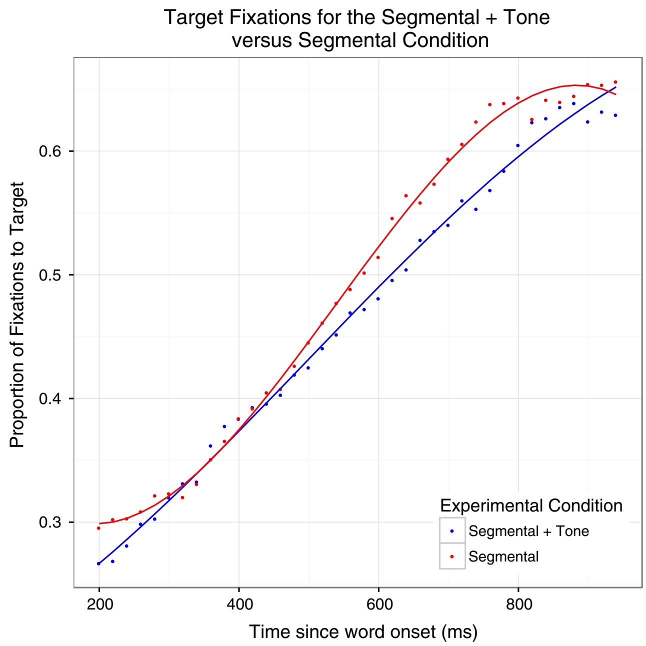


**Fig. 4.** Direct comparison of target fixations in the Segmental + Tone and Segmental conditions. The points represent mean proportions of target fixations at each time step, whereas the lines represent third order growth curve model fits.

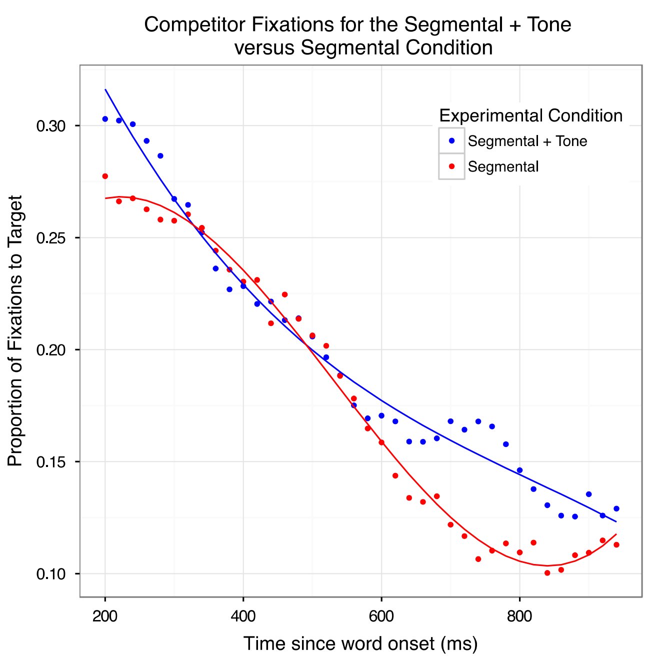


**Fig. 5.** Direct comparison of competitor fixations in the Segmental + Tone and Segmental conditions. The points represent mean proportions of competitor fixations at each time step, whereas the lines represent third order growth curve model fits.

for the Segmental + Tone condition and the baseline intercept for the target fixation model ($r = 0.34$, $p = 0.044$, 95% CI [0.01, 0.60]). In other words, participants who were more proficient in English showed a decreased amount of target fixations in this condition compared to participants who were less proficient in English. The correlation for the competitor fixation model for proficiency was non-significant ($r = 0.089$, $p = 0.605$), as were the correlations with age of acquisition for both the target and competitor fixation models (targets: $r = -0.065$, $p = 0.707$; competitors: $r = -0.066$, $p = 0.701$).

## 4. Discussion

Our study examined language co-activation without input overlap in bilinguals of one tonal language and one non-tonal language, as well as the role of lexical tone in bilingual spoken word recognition through automatic/unconscious translation. We hypothesized that (1) top-down and/or lateral mechanisms are sufficient to induce cross-language lexical competition without any overlap in form; (2) tone is a critical cue in accessing a non-target tonal language during automatic/unconscious translation. To address these hypotheses, we asked Mandarin-English bilinguals to complete a visual world paradigm experiment in which the Mandarin translations of English target items overlapped phonologically in segments and/or tones with competitor items.

First, our findings provide evidence for cross-language competition effects without any phonological overlap between targets and competitors, and suggest implicit access to the Mandarin translations of targets and subsequent phonological activation resulting in lexical competition from the non-target language (see Fig. 7). This language co-activation likely occurred as a result of top-down and/or lateral mechanisms in bilingual language processing, consistent with previous results (e.g., Thierry & Wu, 2007) and adding support from online language measures to this line of evidence, but in a different modality (spoken words) and a different linguistic dimension (lexical tones). There are two possible pathways of activation driving the cross-language lexical competition, as shown in Fig. 7. One possibility is that translation originated at the lexico-semantic level in activating the target non-tonal language by L2 input (e.g., *rain*), while sending feedback to the phonological level in the non-target language through implicit access to translations in L1 (e.g., *yu3*), thus eliciting lexical activation of the non-target language through feedforward activation from phonology to semantics (e.g., 羽 'feather'). Similarly, Thierry and Wu (2007) suggest that access to translation equivalents takes place at a late, perhaps post-lexical processing stage (i.e., during and after word meaning retrieval), as they noted an absence of a priming effect before the N400 window in an ERP experiment testing the repetition effect of covert Chinese characters through translations. This pathway suggests that cross-language activation can be lexico-semantically mediated. Another possibility is that translation was the result of direct translation links from L2 to L1, namely from '*rain*' to '*yu*' in Fig. 7, as suggested in the Revised Hierarchical Model (RHM) by Kroll and Stewart (1994) and Kroll, van Hell, Tokowicz, and Green (2010). The RHM predicts that less

**Table 6**
Growth curve analysis results for direct comparisons of the target and competitor fixation models between the Segmental + Tone and Segmental conditions.

| Component | Targets | | | | Competitors | | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | SE | *t* | *p* | Estimate | SE | *t* | *p* |
| Intercept | 0.022 | 0.014 | 1.57 | 0.117 | −0.018 | 0.011 | −1.69 | 0.092 |
| Linear | 0.048 | 0.085 | 0.567 | 0.571 | −0.040 | 0.065 | −0.622 | 0.534 |
| Quadratic | −0.027 | 0.066 | −0.412 | 0.680 | −0.028 | 0.049 | −0.569 | 0.570 |
| Cubic | −0.088 | 0.040 | −2.20 | 0.028 | 0.089 | 0.038 | 2.33 | 0.020 |

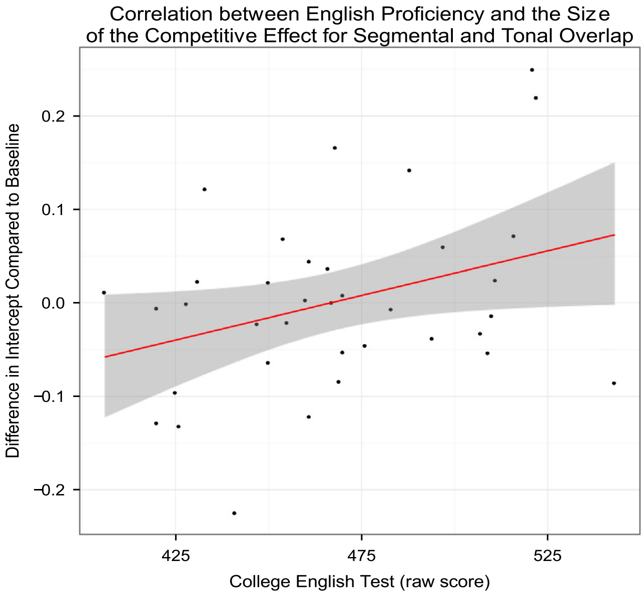


**Fig. 6.** The correlation between L2 English proficiency as measured by CET scores and the size of the difference in intercept between the Segmental + Tone and baseline growth curve models for target fixations. The shaded region represents the 95% confidence interval for the line of best fit.

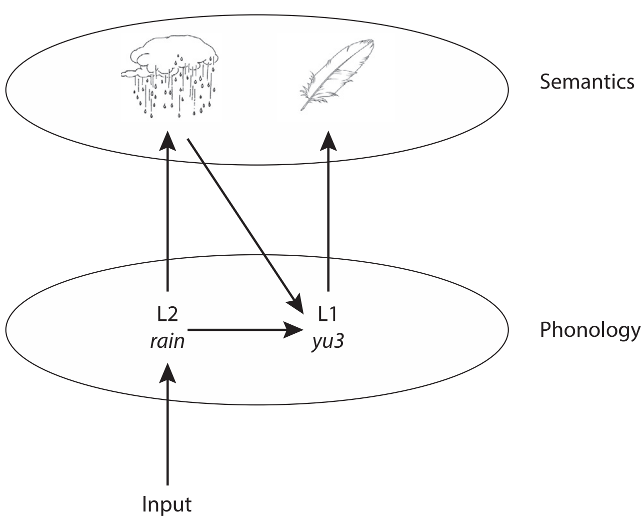


**Fig. 7.** The hypothesized flow of activation in bilingual language processing ('rain' is the target, 'feather' is the competitor; 'yu3' in Mandarin can mean either 'feather' or 'rain'). As shown in this diagram, information flows in a bottom-up fashion from input towards semantics, but can also flow in a top-down fashion from semantics to the phonological layer, and/or via a direct translation route from 'rain' to 'yu3'.

proficient bilingual speakers are more likely to use translation links to access their L2. Namely, recognizing a spoken word in an L2 requires accessing L1 translation equivalents without semantic mediation. The English test scores show that our bilingual participants were spread over a range of L2 proficiency, and it is likely that less proficient bilinguals relied on translation links when processing their L2. This translation via direct L2-L1 excitatory connections without semantics can also result in this type of lexical interference effect and subsequently elicit lexical activation in the non-target language (e.g., 羽 ,'feather'). However, this type of translation/associative mechanism without lexico-semantic mediation is often inconsistent with evidence from the bilingual visual word recognition literature, which has instead supported lexico-semantically driven translation effects, possibly with more proficient bilinguals (e.g., Schoonbaert, Duyck, Brysbaert, & Hartsuiker, 2007; Wang & Forster, 2010, 2014). Even though the lexical competition effects we observed became apparent relatively late in the time course (around 500 ms), we cannot completely rule out the cross-linguistic lateral pathway that can contribute to cross-language lexical competition. Nevertheless, both mechanisms confirmed our first hypothesis: bilingual lexical access to the non-target language does not require input overlap between languages. Future research should investigate and tease apart these two different mechanisms driving the same effect.

Second, our results demonstrate that *only* the Segmental + Tone condition produced significant competition effects when compared to baseline, or when compared directly to the Segmental condition. These differences were apparent for both target and competitor fixations, especially between 500 and 800 ms. These results suggest that both tonal and segmental information are critical in accessing the L1 Mandarin lexical representations when processing L2 English, corroborating the critical nature of both tonal and segmental cues in Mandarin word recognition (e.g., Malins & Joanisse, 2010), and confirming our second hypothesis. The failure to observe competition in the Segmental and Rime + Tone conditions suggests that cross-language lexical competition driven by top-down and/or lateral mechanisms requires an exact match in phonological form for tonal bilinguals. This is in line with the TRACE-T model, which encodes both Mandarin tones and phonemes in the middle layer of representations, such that continuous mapping to relevant representations takes place as both suprasegmental and segmental information become available (Shuai & Malins, 2016).

Third, our results showed a positive correlation between L2 proficiency and the size of the difference between the intercept for the Segmental + Tone condition and the baseline intercept in the target fixation growth curve model. This indicates that the more proficient listeners were in their L2, the more delays they demonstrated in fixations to targets. This result is more consistent with the activation pathway through lexico-semantic mediation rather than lateral translation links, as the lateral mechanism predicts that greater proficiency should reduce the reliance on L1 translations during L2 processing. Plausibly, greater L2 proficiency elicited stronger activation at the lexico-semantic level (Brysbaert & Duyck, 2010; Dijkstra & van Heuven, 2002; Wang, 2013; Wang & Forster, 2010), resulting in stronger activation of L1 translations and more extensive activation being sent back to the phonological level, which was then mapped to lexical representations in the non-target language. This process would thus have given rise to a decreased amount of target fixations in the Segmental + Tone condition in more proficient listeners. This finding suggests that proficiency could be the key contributor in guiding bilingual lexical access (semantic mediation vs. lateral connection) to produce this translation effect.

Our results are consistent with two visual world paradigm experiments studying bimodal bilinguals (Giezen et al., 2015; Shook & Marian, 2012). These experiments offered evidence that spoken words in English induced lexical competition in arrays including competitors that shared different types of phonological relationships with the ASL (American Sign Language) translations of target items. For instance, Shook & Marian (2012) reported lexical activation to the competitor picture 'paper' in a trial with the English target word 'cheese', as 'paper' and 'cheese' share the same location and handshape features and only differ in movement features in ASL. They attributed these cross-language effects to the shared phonology in the non-target language ASL between targets and competitors, due to the activation of ASL translation equivalents for targets. In a similar way, our current results were realized through translation equivalents in the bilingual visual world paradigm. However, our tonal bilingual results also appear to differ from the bimodal bilingual results in the extent to which phonological information in the non-target language needs to be similar between targets and competitors in order to mediate cross-language activation. If the mechanism for cross-language activation is universal across unimodal and bimodal bilinguals, we should have observed competition effects in the Segmental only and Rime + Tone condition, based on the previous findings from bimodal bilinguals. However, our results show that only an exact match between both tones and segments is sufficient to lead to word recognition in L1 Mandarin, which is consistent with the Mandarin spoken word recognition literature. In addition, these effects become apparent relatively early (250–300 ms) in bimodal bilinguals, whereas the effects we observed for unimodal bilinguals were relatively late (500 ms). Therefore, this discrepancy between unimodal and bimodal bilinguals is very likely due to language-specific and modality-specific mechanisms in word recognition: because lexical access in ASL only relies upon visual information, differences may have been observed in the sensitivity of the visual world paradigm to different types of phonological competition.

Our results provide the first and direct evidence of implicit activation of lexical tone through cross-language top-down and/or lateral mechanisms without input overlap between languages, and also offer new insights into current bilingual theories. Current bilingual models are mainly input-driven, and do not seem to specify a top-down mechanism that would allow for lexical access to translation equivalents (e.g., the RHM, Kroll & Stewart, 1994; the BIA+ model, Dijkstra & van Heuven, 2002; the IC model, Green, 1998, the BLINCS model, Shook & Marian, 2012, 2013). Our results suggest that top-down and lateral activation mechanisms are sufficient in and of themselves for bilingual word recognition, without additional contributions resulting from input overlap between a target and non-target language (i.e., L1 and L2). Thus, bilingual models should incorporate this top-down mechanism to allow spreading activation to feed back to the phonological level cross-linguistically; furthermore, our current results suggest that this mechanism should be sensitive to language proficiency. In addition, bilingual models need to incorporate supra-segmental information as a layer of representation for tonal languages.

## 5. Conclusion

In this study, we tested the hypothesis that lexical tones are a critical cue in eliciting cross-language lexical competition in bilingual word recognition even if the target language is non-tonal and phonologically unrelated to the non-target language. To do this, we adopted the visual world paradigm, a highly sensitive and well-established method for studying spoken word recognition in different populations including bilinguals, to record participants' eye-movements while processing their L2 English. We found that when bilinguals are given input in a target non-tonal language (English), this can result in cross-language top-down and/or lateral activation of lexical tones in a non-target tonal language (Mandarin). This behavioral evidence is the first demonstrating implicit lexical tone activation in bilingual spoken word recognition due to implicit access to translation equivalents. These results inform current bilingual theories and computational models by not only providing novel evidence for language co-activation in a top-down and lateral manner without any contribution from the input, but also by offering the first demonstration of the critical role of lexical tones in bilingual lexical access.

## Acknowledgements

We thank Viorica Marian for her contributions to this project and acknowledge the support of grant NICHDR059858 during the development of this study in 2012.

Preliminary findings from this study were presented at the Annual Meeting of the Psychonomic Society 2015, held in Chicago, USA.

## Funding

This project is funded by John Fell OUP (EPD06550) from the University of Oxford to Xin Wang. J. G. M. was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P01HD001994 to Haskins Laboratories. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## Appendix A. Supplementary material

For raw data, please use the following public link: https://osf.io/hxpsw/. Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2017.07.013.

## References

Audacity Development Team (2013). Audacity (version 2.0.3) [computer software]. Available from <http://audacity.sourceforge.net/>.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1995). *The CELEX lexical database (CD-ROM)*. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

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

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of Bilingual Language Processing after Fifteen Years of Service? *Bilingualism: Language and Cognition, 13*(3), 359–371.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLOS-ONE, 5*(6), e10729.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.

Cutler, A., & Chen, H. (1997). Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics, 59*, 165–179.

Dijkstra, T., & van Heuven, WJB. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Giezen, M. R., Blumenfeld, H. K., Shook, A., & Marian, V. (2015). Parallel language activation and inhibitory control in bimodal bilinguals. *Cognition, 141*, 9–25.

Green, David W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(2), 67–81.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge, UK: Cambridge University Press.

Howie, J. (1976). *An acoustic study of Mandarin tones and vowels*. Cambridge, UK: Cambridge University Press.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*, B23–B32.

Ju, M., & Luce, P. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science, 15*(5), 314–318.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149–174.

Kroll, JF., van Hell, JG., Tokowicz, N., & Green, DW. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*(30), 373–381.

Lee, C. (2007). Does horse activate mother? Processing lexical tone in form priming. *Language and Speech, 50*(1), 101–123.

Liu, S., & Samuel, A. G. (2007). The role of Mandarin lexical tones in lexical access under different contextual conditions. *Language and Cognitive Processes, 22*(4), 566–594.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language, 62*(4), 407–420.

Malins, J. G., & Joanisse, M. F. (2012). Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese. *Neuropsychologia, 50*(8), 2032–2043.

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Chapman and Hall/CRC.

R Core Team (2014). R: A language and environment for statistical computing (version 3.1.3) [computer software]. R Foundation for Statistical Computing, Vienna, Austria. Available from <https://www.R-project.org/>.

Schirmer, A., Tang, S., Penney, T. B., Gunter, T. C., & Chen, H. (2005). Brain responses to segmentally and tonally induced semantic violations in Cantonese. *Journal of Cognitive Neuroscience, 17*(1), 1–12.

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2007). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition, 37*(5), 569–586.

Sereno, J. A., & Lee, H. (2015). The contribution of segmental and tonal information in Mandarin spoken word processing. *Language and Speech, 58*(2), 131–151.

Shook, A., & Marian, B. (2012). Bimodal bilinguals co-activate both languages during spoken comprehension. *Cognition, 124*(3), 314–324.

Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition, 16*(2), 304–324.

Shuai, L., & Malins, J. G. (2016). Encoding lexical tones in jTRACE: A simulation of monosyllabic spoken word recognition in Mandarin Chinese. *Behavior Research Methods*, 1–12.

Spivey, M., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*, 281–284.

Taft, M., & Chen, H. (1992). Judging homophony in Chinese: The influence of tones. In H. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 151–172). Oxford, England: North-Holland.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*, 557–580.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of National Academy of Science USA, 104*, 12530–12535.

Tong, Y., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and supra-segmental features in Mandarin Chinese. *Language and Cognitive Processes, 23*(5), 689–708.

Wang, X. & Hui, B., 2017. Perceiving lexical tones when listening to English: evidence from Mandarin-English bilinguals (submitted for publication).

Wang, X. (2013). Language dominance in translation priming: Evidence from balanced and unbalanced Chinese-English bilinguals. *The Quarterly Journal of Experimental Psychology, 66*(4), 727–743.

Wang, X., & Forster, K. I. (2010). Masked translation priming with semantic categorization: Testing the Sense Model. *Bilingualism: Language and Cognition, 13*, 327–340.

Wang, X., & Forster, K. I. (2014). Is translation priming asymmetry due to partial awareness of the prime? *Bilingualism: Language and Cognition, 18*(4), 657–669.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1–25.

Wiener, S., & Turnbull, R. (2016). Constraints of tones, vowels and consonants on lexical selection in Mandarin Chinese. *Language and Speech, 59*(1), 59–82.

Wu, Y. J., Cristino, F., Leek, C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition, 129*, 418–425.

Wu, Y. J., & Thierry, G. (2010). Chinese–English bilinguals reading English hear Chinese. *Journal of Neuroscience, 30*, 7646–7651.

Wu, Y. J., & Thierry, G. (2012). Unconscious translation during incidental foreign language processing. *NeuroImage, 59*, 3468–3473.

Yee, E., & Sedivy, J. (2006). Eye movements reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*, 1–14.

Zhang, T., van Heuven, W. J., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science, 22*(10), 1237–1242.

Zhao, J., Guo, J., Zhou, F., & Shu, H. (2011). Time course of Chinese monosyllabic spoken word recognition: Evidence from ERP analyses. *Neuropsychologia, 49*(7), 1761–1770.