

Chinese-English Speakers' Perception of Pitch in Their Non-Tonal Language: Reinterpreting English as a Tonal-Like Language

Language and Speech

1–21

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0023830919894606

journals.sagepub.com/home/las**Marta Ortega-Llebaria** 

Department of Linguistics, University of Pittsburgh, USA

Zhaohong Wu

Beijing Foreign Studies University, China

Abstract

Changing the F0-contour of English words does not change their lexical meaning. However, it changes the meaning in tonal languages such as Mandarin. Given this important difference and knowing that words in the two languages of a bilingual lexicon interact, the question arises as to how Mandarin-English speakers process pitch in their bilingual lexicon. The few studies that addressed this question showed that Mandarin-English speakers did not perceive pitch in English words as native English speakers did. These studies, however, used English words as stimuli failing to examine nonwords and Mandarin words. Consequently, possible pre-lexical effects and L1 transfer were not ruled out. The present study fills this gap by examining pitch perception in Mandarin and English words and nonwords by Mandarin-English speakers and a group of native English controls. Results showed the tonal experience of Chinese-English speakers modulated their perception of pitch in their non-tonal language at both pre-lexical and lexical levels. In comparison to native English controls, tonal speakers were more sensitive to the acoustic salience of F0-contours in the pre-lexical processing due to top-down feedback. At the lexical level, Mandarin-English speakers organized words in their two languages according to similarity criteria based on both F0 and segmental information, whereas only the segmental information was relevant to the control group. These results in perception together with consistently reported production patterns in previous literature suggest that Mandarin-English speakers process pitch in English as if it was a one-tone language.

Keywords

Chinese-English speakers, pitch, bilingual, lexicon, tonal language

Corresponding author:

Marta Ortega-Llebaria, Department of Linguistics, University of Pittsburgh, Cathedral of Learning Room 2830, Pittsburgh PA 15260, USA.

Email: m.ortega.llebaria@gmail.com

Introduction

Pitch in Mandarin is a reliable correlate of tone and is used to express word meanings, such as [ma] with the rising pitch of tone 2 means “numb” and with the falling pitch of tone 4 means “to scold.” As a result, Mandarin speakers only access the meaning “numb” in “ma2” when they process the consonant, vowel, and the tonal pitch shape of the word. However, in stress-accent languages such as English, the pitch shape of the word is modulated by the interaction of lexical stress with sentence intonation (e.g., Beckman, 1986, and Kember, Choi, Yu, & Cutler and Calhoun, Wollum, & Kruse in this volume for the interaction of pitch accent, syntactic, and information structure in the perception of prominence in English; for similar results in Russian see Luchkina & Cole). Consequently, a word such as “Mary” can be uttered with the rising pitch of a question, such as “Mary?” and the falling pitch of a statement, for example “Mary,” without changing its lexical meaning. Because of that, lexical meaning in English is linked to the consonants and vowels of the word, but not its pitch shape. These cross-language differences imply that Mandarin but not English words have the pitch contour specific to the appropriate tone. Given that in the lexicon of bilingual speakers, words in the two languages interact (e.g., Kroll & Tokowicz, 2005), the question arises as to how Mandarin-English bilinguals represent pitch in their bilingual lexicon.

Recent infant research suggests that Mandarin-English bilingual children process pitch differently from monolinguals at 18 months of age (Singh & Foong, 2012; Burnham, Singh, Mattock, Woo, & Kalasnikova, 2018). Singh and Foong’s (2012) study described the pitch processing changes that Mandarin-English bilingual children underwent during their first year of life. At 7.5 months, they recognized Mandarin and English words when matched in pitch contour. At 9 months, they recognized words in both languages regardless of their pitch shape. Only at 11 months, these children recognized Mandarin words when matched in pitch contour and English words regardless of their pitch patterns revealing more appropriate language-specific strategies to pitch processing. These different strategies of pitch processing may underlie the differences Burnham and colleagues (2018) found between 17-month-old Mandarin-English bilingual children on the one hand, and age-matched Mandarin and English monolinguals on the other. The three groups of children listened to a set of Mandarin and Thai words. Results showed that whereas English children were insensitive to pitch changes during word recognition, Mandarin monolinguals and Mandarin-English bilinguals were not. However, only bilinguals were sensitive to pitch contrasts in Thai, the non-native tonal language, revealing that bilinguals had greater sensitivity to tone interpretation. Altogether, these studies suggest that by 18 months, bilingual children process pitch in their tonal language differently from monolinguals. However, whether they process pitch in English differently from monolingual English speakers remains an open question.

Production data by adult Mandarin speakers of English provide one piece of evidence to answer the above question. When reproducing English prosody, non-target pitch productions constitute one consistent pattern of Mandarin-accented English. For example, Mandarin learners of English tend to avoid deaccentuation in post-focal contexts and de-stressing in compounds so the contrast between “a blackboard” and “a black board” becomes blurred (Chen, 2007; McGory, 1997; Ortega-Llebaria & Colantoni, 2014; Tseng, Su, & Visceglia, 2013; Visceglia & Fodor, 2006; Visceglia, Su, & Tseng, 2012). Yes-no questions in English, and in Spanish, are realized with the ascending pitch shape of L*H%, consisting of the nuclear accent L* and boundary tone H% combination. For example, in the question “Is it Mary?”, the low nuclear pitch accent L* is placed in the stressed syllable “Ma” with a rapidly increasing F0 trajectory in the boundary tone H% on the last syllable “ry”. It was observed that both Mandarin learners of English (McGory, 1997) and Mandarin learners of Spanish (Chen, 2007) tended to keep a pitch peak (H*) in the stressed syllable, producing the ascending LH pitch contour in the adjacent unstressed syllable. Thus, these speakers replaced the

L*H% combination with an H* LH% (see also Todaka, 1990, for similar patterns by Japanese speakers of English). As for peak alignment, Mandarin learners of Spanish tended to produce a pitch peak in the stressed syllable of words in declarative sentences, namely H*, instead of the post-tonic syllable as native speakers do, namely L*H (Chen, 2007). Altogether, these L2 production data strongly suggest that regardless of the L2 intonation requirements of deaccentuation, de-stressing, and peak alignment, Mandarin speakers of English or Spanish represent lexical stress with a pitch peak H* on the stressed syllable giving the words of their non-tonal L2 a rather fixed and tonal-like F0 shape.

Likewise, in environments of language contact, Englishes with a tonal substrate such as Cantonese and Nigerian English exhibit the tendency of avoiding deaccentuation (Gussenhoven, 2014; Gussenhoven & Udofot, 2010; Gut, 2008). Moreover, these Englishes displayed a reduced repertoire of pitch accents in comparison to British English. Gussenhoven (2014) showed that because of this reduced pitch accent inventory, speakers of Cantonese English produced the same sentences and words with less pitch movement than speakers of British English, reinforcing the hypothesis that a tonal language, either as the substrate language in contexts of language contact or as the L1 of L2 English learners, somehow re-shapes the pitch variation of words in intonation languages.

As for perception, only three papers, to our knowledge, addressed the question of whether and how the tonal language of bilingual speakers shapes the pitch perception in their non-tonal language (Ortega-Llebaria, Nemoga, & Presson, 2017; Shook & Marian, 2016; Wang, Wang, & Malins, 2017). These studies used lexical decision tasks with English words that varied in pitch and administered them to adult Chinese-English speakers. In Shook and Marian's (2016) study, Chinese-English bilinguals were asked to listen to an audio English word and choose its correct translation out of two Mandarin written words. Participants were faster when the pitch shape of the English word matched that of the tone in the written Mandarin translation, revealing that these speakers could not avoid processing pitch in English words regardless of the fact that in English, pitch shapes are not linked to lexical meaning. Similarly, Wang, Wang, and Malins (2017) showed that bilingual Chinese speakers of English activated Chinese tone when listening to English words via equivalent translation. For example, when listening to English "rain," they will activate the Mandarin equivalent "雨yu3" ("rain"), and as yu3 is homophonous, "羽yu3" ("feather") will be activated as well. However, competitors whose tones did not overlap with the translations of the English target words, for example, "鱼yu2" ("fish") remained unactivated. Further evidence was obtained by Ortega-Llebaria et al. (2017). They administered a lexical decision task with English words to Mandarin speakers of English, Spanish speakers of English, and native English speakers. Prime-target pairs differed in segment and/or pitch shape. Results showed that only Mandarin speakers of English recognized English words with a falling pitch faster than words with a rising pitch revealing a bias towards falling pitch English words. This bias singled out Mandarin-English bilinguals from non-tonal L2 English speakers and native English speakers supporting Shook and Marian's (2016) and Wang, Wang, and Malins' (2017) results of an L2 pitch processing pattern in English different from that of monolinguals. Moreover, Ortega-Llebaria and colleagues (2017) found some cross-language similarities. Mandarin-English speakers and non-tonal speakers were faster at retrieving target words in full matches (e.g., riceR-riceR; riceF-riceF; where R stands for rising pitch and F for falling pitch) than in F0 mismatches (e.g., riceR-riceF; riceF-riceR) showing that regardless of their tonal background, speakers could not avoid processing pitch during word recognition. It was reasoned that the falling-F0 bias had to take place at a higher processing level—for instance, in the lexicon—than the cross-language pattern because only the former processing pattern singled out tonal speakers, who process pitch at the lexical level. It was also hypothesized that whereas non-tonal speakers processed pitch in English words only at early pre-lexical

level—as for instance, to normalize pitch differences due to gender—Mandarin-English bilinguals kept on processing pitch at latter stages, namely, at the lexical level. However, these three studies examined pitch only in English words, excluding nonwords and Mandarin words. Consequently, possible pre-lexical processing and L1 transfer effects were not ruled out, failing to provide direct evidence for the hypothesis that Mandarin-English speakers processed pitch in English words in a tone-like manner in their bilingual lexicon.

The goal of this study was to provide this missing evidence by adding Mandarin words and nonwords to the previous Ortega-Llebaria et al. (2017) lexical decision task. To that end, non-word processing was also examined in the English lexical decision task, and a new but equivalent task was created in Mandarin. A group of Mandarin speakers of English were asked to perform both tasks and an English monolingual control group performed the English task. Although comparing pitch perception in English words by the Mandarin-English speakers and native English controls constituted a replication of the previous study, additional examination of nonwords provided the missing evidence to test whether the patterns that singled out Mandarin-English speakers took place at the pre-lexical level. Moreover, the comparison between Mandarin and English words and nonwords processing by Mandarin-English bilinguals allowed us to test possible L1 transfer effects.

2 Methodology

2.1 Participants

Two groups of speakers participated in the study. The experimental group was constituted by 40 native speakers of Mandarin living and studying an undergraduate degree at Nankai University in Tianjin, China at the time of testing. They learned Mandarin at home (34 participants) or at around age 5 or 6 when starting school (six participants). After that they used it regularly at school and with friends. The 40 participants had started learning English since they were about 6 or 7 years old in primary school. They continued learning English at Nankai University as a subject taught once or twice per week. However, none majored in English and none had lived in an English-speaking country. Their self-rating on their English reading, writing, listening, and speaking abilities indicated that their English proficiency was intermediate. Each received \$10 as compensation for their participation.

In total, 35 native English speakers constituted the control group. They were undergraduate students living and studying at University of Pittsburgh at the time of testing. Most English participants studied a second language in high school and at the university. However, none studied or were fluent in a tonal language.

No participants reported having been diagnosed with any speech and hearing problems.

2.2 Materials and recordings

Mirroring the lexical decision task used in Ortega-Llebaria et al. (2017), we prepared materials for a lexical decision task in English and an equivalent one in Mandarin. Words (e.g., *mice*, *tree*) and nonwords (e.g., *kice*, *pree*) in each language were read by a native speaker. For English, the native speaker was a 40-year-old male speaker of Standard American English. For Mandarin, the 23-year-old male was a native speaker of Mandarin from Beijing. Speakers were recorded directly into a computer with Praat software at 44,100 Hz sampling rate with a Sennheiser Evolution e845 microphone in the quiet room of the recording studio at the Language Media Center at University of Pittsburgh. Both speakers had very clear diction and were instructed to

read at a normal speaking rate, pausing after each item. The authors were present during the recordings. Special attention was paid to the intonation of words so each word was clearly spoken with a falling and a rising pitch contour. All words were monosyllabic and respected the syllabic constraints of the language. The carrying sentences “I say ____” and “Did you say ____?” and their Mandarin equivalents “wo3 shuo1,” “I say,” and “ni3 shuo1 de shi4 ____ma?” “Did you say ____?” were used to elicit the production of these two pitch contours. If the authors considered an item was not sufficiently clear, the native speakers were asked to repeat it, and when needed, the authors provided an oral example for imitation.

The monosyllabic words and nonwords in both languages differed by one consonant, for example *mice-kice*, *tree-pree* in English and *tao-fao*, *ma-ka* in Mandarin. Each word and nonword was read with two F0 contours, with the rising intonation of a question and the falling intonation of a statement in English, such as *miceR*, *kiceR*, *miceF*, *kiceF*, and in Mandarin, with the rising pitch of tone 2 and the falling pitch of tone 4. Given the high degree of homophony in Mandarin, the second author, a native Mandarin speaker, ensured nonwords were clearly understood as nonwords by checking the nonwords had no entry in the Xinhua Zidian—or Xinhua Dictionary—published by the Commercial Press in China. Durations of words and nonwords were measured. Items with a rising F0 tended to have longer durations than those with falling F0. This difference was controlled for in the statistical model as explained in Section 2.4.

Recorded words were excised using Praat (Boersma & Weenink, 2018) from their respective carrying sentences at zero crossings to avoid clicks. After levelling them for intensity at 70 dB, they were concatenated into prime-target pairs with a 250 ms of silence between the two words. These pairs were classified into five conditions according to the prime-target similarity. In the Full Match condition (FM), prime and target words shared the same segments and F0 contour, such as *miceR-miceR* in English or *ma2-ma2* in Mandarin, whereas in the Full Mismatch condition (FMisM), there were no phonemes in common between prime and target and the two words had opposite F0 contours, for example, *goldF-miceR* in English or *bao4[bau]-ma2[ma]* in Mandarin. These two conditions served as baselines for the partial mismatched conditions, where prime-target differences elicited the effect of segmental (Mismatched Segments, MMS) and pitch contour differences (Mismatched F0, MMF0). In particular, the mismatched segments in the MMS condition, which were pronounced as different phonemes, resided in the onset position in Mandarin, such as *pa2-ma2*, whereas in English appeared in either onset or coda, for example, *riceR-miceR*, *plateF-planeF*. In the MMF0, primes and targets differed in the pitch contour while sharing the same segments, such as *miceF-miceR* in English or *ma4-ma2* in Mandarin. Finally, a fifth condition, the MMS and F0 (MMSF0) was added. Prime-target pairs in the MMSF0 differed in one segment and in the F0 contour, for example, *riceR-miceF* in English or *pa4-ma2* in Mandarin, to elicit segment and pitch effect within the same trial.

In English, the triplets formed by the prime in MMS, target, and the prime in FMisM (e.g., rice, mice, gold, see Appendix 1 for all triplets) had similar frequencies and numbers of neighbors. Across triplets, the log of frequency of a word as reported by the HAL study (Log_Freq_HAL) ranged from 8.3 to 10.2 and the number of phonological neighbors from 11–50, based on the English Lexicon Project (Balota et al., 2007). Because in Mandarin tone conditions word frequencies, it was ensured that all the primes for the same target—for example, for target *ma2*, the primes are *ma2* in FM, *bao4* in FMisM, *pa2* in MMS, *ma4* in MMF0, and *pa4* in MMSF0—had the same frequency and neighborhood density. Thus, in Mandarin, the primes and targets formed quintets, and within each quintet, words had the same frequency and neighborhood density. Across quintets, frequency per million ranged from 0.00 to 0.16 and neighborhood density from 8 to 26 based on the Database of Mandarin Neighborhood Statistics (Neergaard, Xu, & Huang, 2016).

2.3 Lexical decision tasks

As shown in Appendix 1, the English task contained 312 prime-target pairs (144 word targets, 168 nonword targets) and the Mandarin task contained 320 pairs, of which 260 contained tones 2 and 4 (141 word targets, 119 nonword targets) and the remaining were fillers containing tones 1 and 3. The English task contained 12 groups (see lines 1–12 in Appendix 1: six groups of onset alternation and six groups of coda alternation for MMS), and the Mandarin task contained 10 groups (see lines 1–10 in Appendix 1: all onset alternations for MMS). Word targets were presented in all five priming conditions (FM, FMisM, MMS, MMF0, MMSF0) and corresponding nonword targets were presented after the same primes as the word targets. In addition, FM and MMF0 condition primes were included for the nonword targets. To avoid a situation of nonword priming having only nonword targets, we included a nonword prime-word target condition. Therefore, in each group, there were 26 prime-target pairs. The SOA between prime and target was 250 ms and the inter-trial interval in both the English task and the Chinese task was 1,000 ms.

2.3.1 Procedure. The Mandarin participants listened to both the English and the Mandarin priming lexical decision tasks via Sennheiser headphones connected to the computers in the Psychology Laboratory at Nankai University, Tianjin, China. A random half of Mandarin participants completed the experiments first in English and second in Mandarin and the other half completed the experiments in the opposite order. The English participants listened to the English priming lexical decision task via Sennheiser headphones connected to the computers in the Psychology Laboratory at University of Pittsburgh.

Participants were asked to listen to the prime-target pairs in each trial and decide if the target was a real English/Mandarin word, by pressing "j" if they heard a real English/Mandarin word and "f" if they heard a nonword. After reading the instructions on the computer screen, participants had a practice session that consisted of eight pairs from different conditions (four word targets, four nonword targets for both English and Mandarin). After the practice, they continued to the main experiment. The real experiment started with a warm-up of 10 pairs from different conditions (five word targets and five nonword targets for English; six word targets, and four nonword targets for Mandarin) to give participants time to get acquainted with the task and the keypad. Those 18 pairs (practice and warm up) were not included in the results. The critical trials were those after the warm up.

Participants were given a short break after every 100 stimuli, for a total of three breaks. The English and Mandarin experiments took approximately 22–24 minutes each for the Mandarin-English speakers, whereas it took approximately 18–20 minutes for the native English speakers. Participants were asked to complete a language history questionnaire at the end of the study.

2.4 Data analysis

Trials with negative reaction times (RT, i.e., when participants responded before the end of the target presentation) and those above 2500 ms were removed from the data. The log-transformed reaction time (LRT) data were analyzed using linear mixed-effects models in *R* (R Core Team, 2014) and the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) on the correctly responded trials by Mandarin-English speakers (Mandarin task: 81.93% correctly identified words, 82.49% nonwords; English task: 80.27% words and 63.06% nonwords) and by the native English controls (English task: words: 91.10%, nonwords: 90.46%). The high percentage of correctly identified words and nonwords in the Mandarin task indicates that despite the high degree of homophony in Mandarin words, Mandarin-English speakers did not confuse nonwords as possible words.

Accuracy (ACC) data were analyzed using generalized (logit) linear mixed-effects models with the same statistical package.

The fixed effects of interest included Priming Condition, Pitch Shape, Lexicality, and their two-way and three-way interactions. Target Duration was included in the models as a covariate to partial out effects of target duration on reaction time or accuracy. Frequency and Neighborhood Density were included in the models if model comparisons showed a need to include them. The random effects included the random intercepts for Subject and Group (i.e., which item the prime-target pairs was from, see Appendix 1 for the list of prime-target pairs), and random slopes were included in the model as well if model comparisons showed a need to include them.

To foreshadow the results, lexicality was found to interact with either Priming Condition or Pitch Shape for both the Mandarin and English tasks. Therefore, separate analyses were conducted for the word data and the nonword data for the Mandarin-English participants and native English speakers. The specific formulae used in each model are listed in Appendix 2.

3 Results

3.1 Lexicality effects

Results showed that Mandarin words were identified faster than Mandarin nonwords by Mandarin-English speakers ($F = 53.84, p < 0.001$), but not significantly more accurately (Chi square, $\chi^2 = 0.15, p > 0.05$). English words were identified faster ($F = 45.68, p < 0.001$) and significantly more accurately ($\chi^2 = 252.70, p < 0.001$) than English nonwords by Mandarin-English speakers. In contrast, there was no significant difference in reaction time ($F = .35, p > 0.05$) or accuracy ($\chi^2 = 1.48, p > 0.05$) for words and nonwords by native English controls. Therefore, Lexicality had a significant effect in Mandarin-English speakers in both the Mandarin and the English lexical decision tasks. The three-way interaction among Condition, Pitch Shape, and Lexicality for native Mandarin speakers in the Mandarin experiment did not reach statistical significance for LRT ($F = 2.06, p = 0.08$) or for ACC ($\chi^2 = 1.08, p > 0.05$), whereas the three-way interaction in the English experiment for ACC did ($\chi^2 = 35.31, p < 0.001$). For Native English speakers, however, there was no significant three-way interaction among Condition, Pitch Shape, and Lexicality in the English experiment for LRT ($F = 1.01, p > 0.05$) or for ACC ($\chi^2 = 6.50, p > 0.05$). Moreover, for the Chinese task for Mandarin-English speakers, for both reaction time and accuracy, there were two-way interactions between Priming Condition and Lexicality ($F = 4.23, p < 0.01; \chi^2 = 28.58, p < 0.001$) and between Pitch Shape and Lexicality ($F = 6.67, p < 0.01; \chi^2 = 46.48, p < 0.001$). Similarly, for the English task for Mandarin-English speakers, for reaction time, there was an interaction between Pitch Shape and Lexicality ($F = 28.38, p < 0.001$), and for accuracy, there were two way interactions between Priming Condition and Lexicality ($\chi^2 = 21.38, p < 0.001$) and between Pitch Shape and Lexicality ($\chi^2 = 280.74, p < 0.001$). For native English speakers, for reaction time, there was a significant interaction between Priming Condition and Lexicality ($F = 3.05, p < 0.05$).

As a result, separate models were run for words and nonwords for both the Mandarin task and the English task.

3.2 Effects of Priming Condition and Pitch Shape in words and in nonwords

Table 1 and Figure 1 summarize the results from models analyzing log-transformed RT in words and nonwords separately for Chinese-English participants. As explained in Section 2.4, the variable Target Duration controlled for the fact that target words with rising F0 had longer durations

Table 1. Log-Transformed Reaction Times for Native Chinese and Native English Speakers.

Chinese speaker		<i>F</i>	Pr(> <i>F</i>)	
LRT Chinese word	PrimingCondition	48.63	0.00	***
	PitchShape	6.22	0.01	*
	PrimingCondition:PitchShape	1.39	0.24	
LRT Chinese nonword	PrimingCondition	12.88	0.00	***
	PitchShape	0.00	0.94	
	PrimingCondition:PitchShape	4.08	0.00	**
LRT English word	PrimingCondition	31.86	0.00	***
	PitchShape	23.36	0.00	***
	PrimingCondition:PitchShape	2.94	0.02	*
LRT English nonword	PrimingCondition	14.92	0.00	***
	PitchShape	6.59	0.01	*
	PrimingCondition:PitchShape	1.18	0.32	
English speaker				
LRT English overall	PrimingCondition	17.49	0.00	***
	PitchShape	0.04	0.85	
	PrimingCondition:PitchShape	1.67	0.15	

Statistically significant differences between falling and rising targets in each condition are marked with * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

than their falling F0 counterparts, allowing to partial out the effect of duration in the comparison of their reaction times. The duration difference between rising and falling targets was 58.20 ms for Mandarin words and 67.27 ms for Mandarin nonwords, and 39.14 ms for English words and 4.29 ms for English nonwords. Based on Table 1 and Figure 1, two main patterns stand out. First, Priming Condition modulates the reaction times of all participants similarly with a $p < 0.001$, regardless of their language background. Second, Pitch Shape modulates log-transformed RT in Mandarin-English speakers, but not in native English controls. This Pitch Shape modulation observed only in Mandarin-English speakers was statistically significant in both of their languages as a main effect (Mandarin words: $F = 6.22$, $p < 0.05$; English words: $F = 23.36$, $p < 0.001$; English nonwords: $F = 6.59$, $p < 0.05$) and as an interaction with Priming Condition (Mandarin nonwords: $F = 4.08$, $p < 0.01$; English words: $F = 2.94$, $p < 0.05$). These patterns are explained in detail in Sections 3.2.1 and 3.2.2, respectively.

Table 2 and Figure 2 summarize the results for accuracy. Basically, for Mandarin-English speakers, there was a significant falling advantage in all priming conditions for Mandarin words, English words, and Mandarin nonwords. However, in English nonwords, there was a significant rising advantage in all priming conditions. For native English speakers, there was no effect of accuracy on words and nonwords except a falling advantage for words in the mismatched F0 condition. These patterns are explained in detail in Section 3.2.1 Priming Condition, and 3.2.2 Pitch Shape.

3.2.1 Priming Condition. Priming Condition was significant across speakers and languages for log-transformed reaction times at $p < 0.001$ (Table 1), and across the languages of Mandarin-English speakers for accuracy at $p < 0.001$ (Table 2), indicating the degree of similarity between primes and targets modulated lexical access. The priming effects for each condition relative to the full match and the full mismatch are detailed in Appendix 3 and are illustrated in Figure 1. Targets in fully matched prime-target pairs (FM in Figure 1) were retrieved faster than in partial mismatches (MMS, MMF0, MMSF0), and in turn, those were retrieved faster than targets in full mismatches

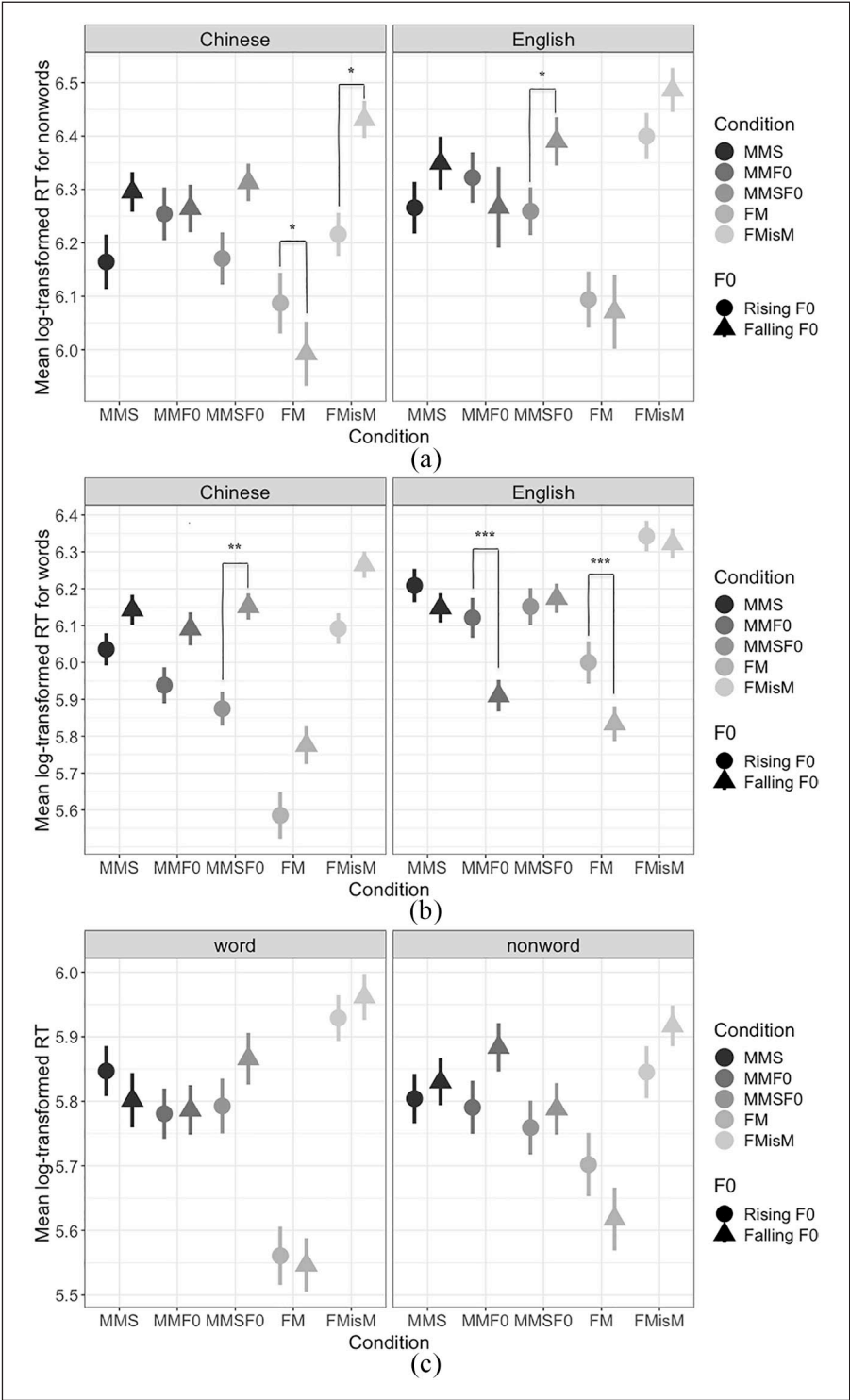


Figure 1. (Continued)

Figure 1. Mean log-transformed reaction times to target identification with falling (triangles) and rising (circles) pitch in each priming condition. Priming conditions are from left to right: MMS: mismatched segments (riceR-miceR), MMF0: mismatched F0 (miceR-miceF); MMSF0: mismatched segments and F0 (riceR-miceF); FM: full matches (miceR-miceR); FMisM: full mismatches (goldR-miceF). (a) refers to the Chinese and English nonwords identified by Chinese-English speakers, (b) to the Chinese and English words identified by Chinese-English speakers, and (c) by the English words and nonwords identified by native English speakers. Statistically significant differences between falling and rising targets in each condition are marked with * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

Table 2. Accuracy Results for Native Chinese and Native English Speakers.

Chinese speaker		χ^2	Pr(> χ^2)	
ACC Chinese word	PrimingCondition	22.68	0.00	***
	PitchShape	22.73	0.00	***
	PrimingCondition:PitchShape	0.93	0.92	
ACC Chinese nonword	PrimingCondition	15.75	0.00	**
	PitchShape	29.28	0.00	***
	PrimingCondition:PitchShape	0.96	0.92	
ACC English word	PrimingCondition	21.45	0.00	***
	PitchShape	180.15	0.00	***
	PrimingCondition:PitchShape	26.08	0.00	***
ACC English nonword	PrimingCondition	29.11	0.00	***
	PitchShape	105.97	0.00	***
	PrimingCondition:PitchShape	15.10	0.00	**
English speaker				
ACC English overall	PrimingCondition	1.37	0.85	
	PitchShape	1.72	0.19	
	PrimingCondition:PitchShape	2.33	0.68	

Statistically significant differences between falling and rising targets in each condition are marked with * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).
 χ^2 : Chi square.

(FMisM). These patterns appeared in both words and nonwords and in the two languages of the Mandarin-English speakers (Tables 1 and 2 of Appendix 3 for F and p values), and in the words and nonwords of native English speakers (Tables 3 and 4 of Appendix 3). For example, Table 1 shows that Mandarin-English speakers generally identified the fully matched targets and the partial mismatched targets faster than the fully mismatched targets in words and nonwords and in their two languages. Table 2 shows similar patterns in relation to Full Matches.

It is worth noting that both Mandarin-English and native English speakers retrieved F0 mismatches (MMF0) more slowly than fully matched pairs in both words and nonwords (see Table 2 for Chinese speakers and Table 4 for English speakers, Appendix 3) suggesting that tonal and non-tonal speakers could not ignore pitch in either English words and nonwords despite the fact that pitch variation in English is not relevant to lexical meaning, and therefore, it has previously been considered irrelevant for lexical access in English words (e.g., Malins & Joanisse, 2012, and references therein).

With regards to accuracy, Tables 5 and 6 of Appendix 3 show for both Mandarin and English words, Mandarin speakers generally identified FM targets more accurately than FMisM targets, but they identified the partial mismatches MMS, MMF0, and MMSF0 significantly less

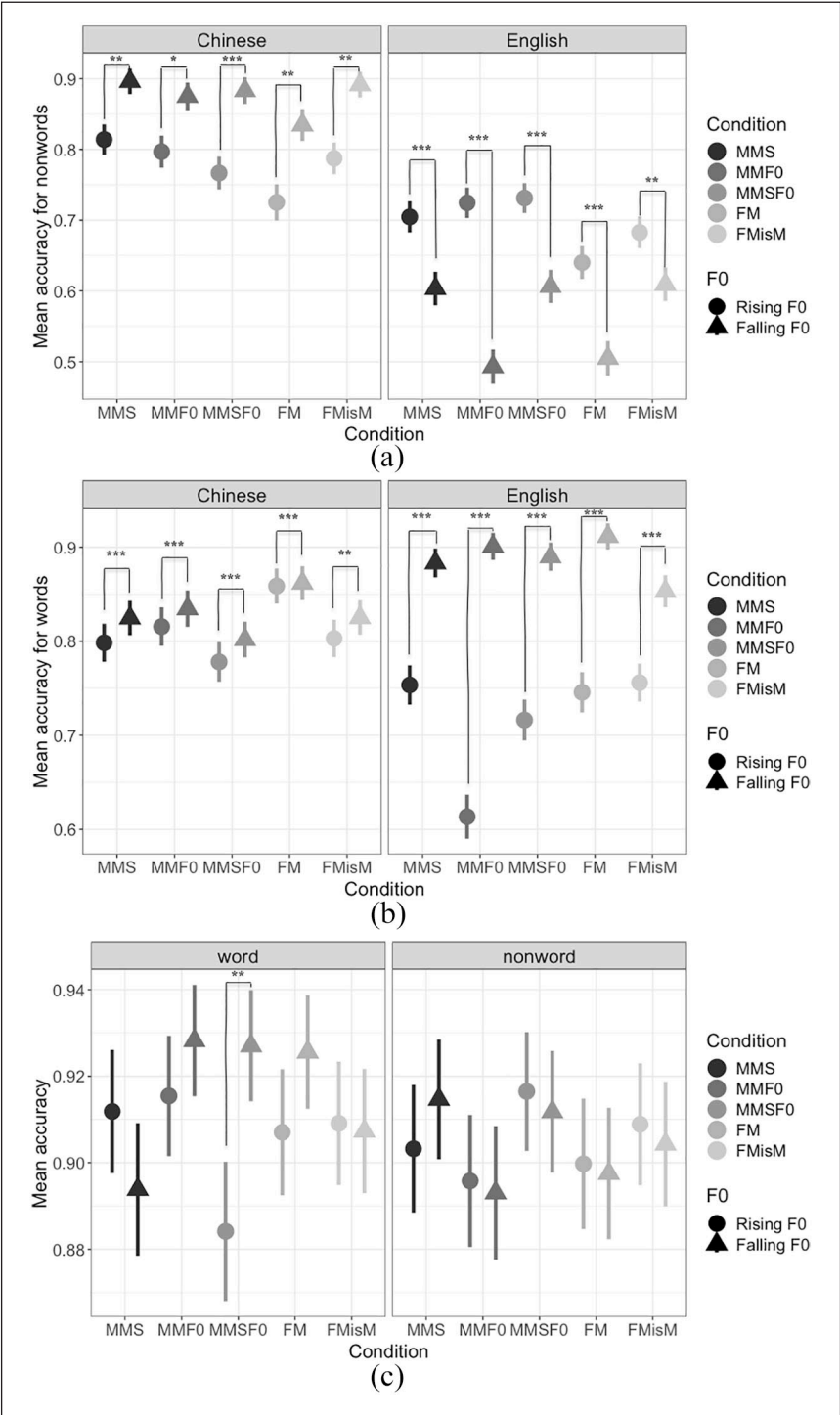


Figure 2. (Continued)

Figure 2. Mean accuracy scores to target identification with falling (triangles) and rising (circles) pitch in each priming condition. Priming conditions are from left to right: MMS: mismatched segments (riceR-miceR); MMF0: mismatched F0 (miceR-miceF); MMSF0: mismatched segments and F0 (riceR-miceF); FM: full matches (miceR-miceR); FMisM: full mismatches (goldR-miceF). (a) refers to the Chinese and English nonwords identified by Chinese-English speakers, (b) to the Chinese and English words identified by Chinese-English speakers, and (c) by the English words and nonwords identified by native English speakers. Statistically significant differences between falling and rising targets in each condition are marked with * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

Table 3. Reaction Time Mean Differences Between Targets with Falling and rising F0 (Falling-Rising).

Participant	Chinese-English speakers				Native English speakers	
Task	Chinese		English		English	
Stimuli	Word	Nonword	Word	Nonword	Word	Nonword
FM	42.812	-45.662	-104.456	15.942	-11.854	-29.36
FMisM	75.039	120.898	-13.424	46.402	18.669	3.994
MMS	59.451	32.813	-45.523	46.284	-4.642	7.096
MMF0	70.618	-12.027	-141.306	-3.131	-9.752	23.085
MMSF0	111.532	30.503	-3.633	62.587	33.296	-7.068
Average	71.8904	25.305	-61.6684	33.6168	5.1434	-0.4506

MMS: mismatched segments; MMF0: mismatched F0; MMSF0: mismatched segments and F0; FM: full matches; FMisM: full mismatches.

accurately than the FM targets. For nonwords, FM is less accurate than FFM. Thus, the more similar the prime-target word pairs were, the more accurately they were perceived. In contrast, there were no significant effects for native English speakers (Tables 7 and 8, Appendix 3). For both words and nonwords, native English speakers generally did not identify FM targets more accurately than FMisM targets, and they did not identify the partial mismatches MMS, MMF0, and MMSF0 more accurately than the FM targets.

To summarize, in Mandarin-English speakers, the degree of phonetic similarity between prime and targets modulated the reaction times and accuracy of target identification in a comparable manner. The stronger the phonetic similarity between prime and target, the faster and more accurately the target was identified. However, for native English speakers this pattern was significant only in reaction times. Interestingly, both tonal and non-tonal speakers retrieved FM faster than MMF0 not only in words, as in the previous study (Ortega-Llebaria et al., 2017), but also in nonwords, suggesting that independently of their tonal background, participants could not ignore F0 variation in English targets despite that pitch in English does not convey lexical meaning.

3.2.2 Pitch Shape. Although the effects of Priming Condition highlighted similarities across speakers and languages, the effects of Pitch Shape showed clear differences between Mandarin-English speakers and the English controls on the one hand, and between the two languages of the Mandarin-English speakers on the other. As for the former, Mandarin-English speakers showed larger reaction time differences between rising and falling F0 targets than native English controls. For example, Table 3 shows the reaction times' mean differences between these targets. Mandarin-English speakers identified the English FM targets with a rising pitch, for example, miceR-miceR, 104.46 ms slower than with a falling pitch, such as miceF-miceF. However, this difference

reduced to 11.85 ms when the same words were identified by native English speakers, suggesting pitch shape had a stronger effect in Mandarin-English speakers than in native English speakers. The results in Table 1 corroborate the statistical significance of these differences in log-transformed RTs. The significant effects of Pitch Shape as a main factor obtained in Mandarin-English speakers (English words: $F = 23.36$, $p < 0.001$; English nonwords: $F = 6.59$, $p < 0.01$) contrasted with the non-significant results obtained in native English speakers ($F = 0.04$, $p > 0.05$). Similarly, pitch shape affects accuracy in Mandarin-English speakers (Table 2, English words: $\chi^2 = 180.15$, $p < 0.001$; English nonwords $\chi^2 = 105.97$, $p < 0.001$) but not in native English speakers ($\chi^2 = 1.72$, $p > 0.05$). Altogether, results from reaction times and accuracy indicate that pitch shape becomes relevant to lexical processing only to Mandarin-English speakers and not to native English speakers.

Differences in Pitch Shape also took place between the two languages of Mandarin-English speakers. As illustrated in Table 3, these speakers retrieved Mandarin words with rising pitch faster than their falling pitch counterparts ($F = 6.22$, $p < 0.05$). However, in English, they retrieved faster words with falling than with rising pitch ($F = 23.36$, $p < 0.001$). This falling pitch advantage of English words was modulated by condition (Pitch Shape \times Condition: $F = 2.94$, $p < 0.05$). The priming effects obtained in the mixed-effects models (see Appendix 4) showed this interaction was related to the Full Match and Mismatched F0 priming conditions. In these conditions, falling pitch targets were perceived significantly faster than their corresponding rising target. It is worth noting that out of the five priming conditions, FM and MMF0 were the only ones where prime-target pairs had identical segments. As for nonwords, rising pitch targets were in general retrieved faster and reached statistical significance in English ($F = 6.59$, $p < 0.05$) but not in Mandarin ($F = 0.00$, $p > 0.05$).

With regard to accuracy, Table 2 shows the effect of Pitch Shape was significant only in Mandarin-English speakers, and Figure 2 indicates that in the Chinese task, these speakers perceived rising targets less accurately. However, results are mixed in the English task, where these participants perceived falling targets more accurately in English words, and rising targets in English nonwords.

To summarize, Pitch Shape modulated the reaction times and accuracy of Chinese-English speakers, but not of native English speakers. For Chinese speakers, targets with a rising pitch were retrieved faster and perceived less accurately in Chinese words and nonwords. Rising pitch targets were also retrieved faster and more accurately in English nonwords. It was only in English words that targets with falling pitch were retrieved faster and more accurately than rising pitch targets, and this effect was modulated by condition. It was in FM and MMF0, the two priming conditions with identical segments in prime and targets, that Mandarin-English speakers retrieved falling pitch targets significantly faster than their rising counterparts.

4 Discussion

4.1 Replicating results: Pitch in English words

Results from Mandarin-English speakers and native English speakers listening to pitch variations in English words in this new experiment faithfully replicated the results obtained in our previous study (Ortega-Llebaria et al., 2017). On the one hand, in both studies there were consistent similarities between speakers. Reaction times in fully matched prime-target pairs were faster than in partial mismatches (mismatch in Segment, mismatch in F0, mismatch in Segment and F0). In turn, those were faster than fully mismatched pairs, showing the phonetic similarity between primes and targets facilitated word access. Moreover, in this experiment, Chinese-English speakers not only

identified more similar prime-target pairs faster but also more accurately. Of particular interest was that speakers from both language groups obtained faster reaction times in fully matched pairs than in pairs with mismatched F0 (MMF0) showing that, regardless of their tonal background, all participants processed pitch despite the fact that English is a non-tonal language. Altogether, these results were interpreted as evidence that similarity computations between primes and targets included both segments and F0 independently of the tonal status of the language being processed and of the tonal background of the listeners. Because these differences took place across language groups, they were attributed a pre-lexical status and explained by a common mechanism of integral perception by which Mandarin and English speakers could not ignore pitch when processing segmental contrasts and vice versa (Lee & Nusbaum, 1993; Miller, 1978; Repp & Lin, 1990). It was proposed that this mechanism would form part of the shared feature detectors proposed in bilingual models of spoken word recognition such as BLINCS (Shook & Marian, 2013) or BIMOLA (Grosjean, 1998; Lewy, 2008) by which pitch variation in the signal was processed and sent up to progressively higher and more abstract pitch representations. At the feature level, however, pitch abstraction is minimal mostly reproducing the F0 variation present in the speech signal.

In contrast, differences between speakers' groups arose when identifying pitch contour differences in English words. In results from both studies, namely, this one and Ortega-Llebaria et al. (2017), native speakers of English obtained non-significant scores for Pitch Shape, indicating that regardless of the fact that the English word had the rising pitch of a question or the falling pitch of a statement, this pitch variation did not affect the lexical access of native English speakers. In contrast, Pitch Shape reached significant results for Mandarin-English speakers in both studies. For them, a target English word with a falling pitch was retrieved faster than its rising F0 counterpart especially in priming conditions where prime and target words shared the same segmental information, that is, Full Match and Mismatch in F0. Altogether, these results showed that Mandarin-English speakers, but not native English speakers, continued processing pitch at the lexical level. Consequently, English words, like Mandarin words, were stored with a pitch shape in the bilingual lexicon of Mandarin-English speakers. For these speakers, an English word with a falling pitch shape was a better representation than the same word with a rising pitch, probably because words in citation form are produced with the H*-L% falling intonation of statements (e.g., Pierrehumbert, 1980). Moreover, the majority of English disyllabic content words have a trochaic (strong/weak) stress pattern (Cutler & Carter, 1987), which confers them a falling pitch pattern to which infants, adults, and L2 learners are highly sensitive (e.g., Juszyc, Cutler, & Redantz, 1993; Cutler & Norris, 1988; Tremblay, 2008).

However, as discussed in the introduction, the pre-lexical and lexical status of the above pitch processing patterns was based only on the fact that pre-lexical patterns took place across language groups and lexical patterns only in the Mandarin-English group. Despite these results being replicated in the present study giving them consistency, the criteria to classify these pitch processing patterns as lexical or pre-lexical only on the basis of cross-language differences and similarities, although a revealing first step, needs to be backed up with more data. In particular, these patterns need to be tested in nonwords to show their pre-lexical status, and in the two languages of bilinguals, namely, Mandarin and English, to consider L1 transfer, which is dealt with in the next section.

4.2 New results: Non-words and pre-lexical processes

As explained in Section 4.1, Priming Condition modulated the RT in words so that the more similar primes and targets were with regard to segmental and F0 information, the faster the target was identified. In this experiment, the same modulating effect of Priming Condition consistently

appeared in nonwords, providing the missing evidence needed to confirm this cost of processing pitch and segmental differences takes place at the pre-lexical level cross-linguistically.

In addition to this similar pre-lexical processing of pitch and segmental differences across languages, nonword results showed as well a cross-language difference, namely, Chinese-English speakers identified faster English nonwords with a rising pitch than their corresponding falling pitch counterparts (e.g., see means in Table 3, and the English nonwords in Figure 1(a)). This rising F0 advantage, however, does not appear when native English speakers listen to the same English nonwords (see Table 3 and nonwords in Figure 1(c)). This cross-linguistic difference can be accounted for by (a) the greater acoustic saliency of rising pitch contours, and by (b) the modulation of acoustic processing by language experience via feedback from higher- to low-level processing. Evidence for the acoustic saliency of rising pitch comes from behavioral (Wayland, Zhu, & Kaan, 2015) and neurophysiological studies (Krishnan, Xu, Gandour, & Cariani, 2004; Krishnan & Parkinson, 2000). Wayland and colleagues (2015) found that both tonal and non-tonal speakers discriminated better pairs of tone combinations that contained a rising pitch, for example, rising T2 versus falling T4, than combinations that did not contained it, for example, high T1 versus falling T4. As for Event-Related P studies, Krishnan and Parkinson (2000) found that Frequency Following Responses (FFR), the brainwaves with subcortical origin that encode the acoustic waveform with minimal abstraction, reproduced better rising than falling tonal sweeps showing a greater neural synchrony with rising tones. Krishnan and colleagues replicated these results with words (Krishnan et al., 2004), where Mandarin words yi2 and yi3 obtained more robust FFRs than those for yi1 and yi4, corroborating the acoustic saliency of rising tonal shapes. Altogether, these studies suggest pitch is processed cross-linguistically at a very early subcortical and pre-lexical stage, where rising pitch, due to its acoustic saliency, is initially encoded via robust signals by a mechanism similar to our proposed feature detectors that encode pitch variation very early in processing with a minimum of abstraction.

The fact that only Chinese-English speakers, but not native English speakers, show this rising pitch advantage when listening to the same nonwords is puzzling if we consider that all speakers, tonal and non-tonal, detect pitch differences at pre-lexical stage via feature detectors that encode the acoustic saliency of rising pitch with robust signals. This apparent contradiction can be explained by the top-down modulation that language experience exerts on low-level acoustic encoding. Neurophysiological research showed that FFR, the brainwaves with subcortical origin that encode the acoustic waveform, are modulated by long-term input, demonstrating that the low-level processing of acoustic input is mediated by higher-level knowledge (Intartaglia, et al., 2016; Krishnan et al., 2004, Krishnan, Gandour, & Suresh, 2014; Krishnan, Gandour, Ananthakrishnan, & Vijayaraghava, 2014; but see also Coffey, Herholz, Chepesiuk, Baillet, & Zatorre, 2016). For example, Intartaglia and colleagues (e.g., Intartaglia et al., 2016) showed that English speakers and French speakers displayed stronger FFR for syllables in their respective native languages, revealing that long-term memory phonemic representations of the speakers' L1 had an effect in low-level encoding of acoustic detail. Similarly, the tonal languages' requirement to preserve tonal information in the lexicon may increase Mandarin-English speakers' sensitivity to F0 at pre-lexical stages. The shared feature detectors of Chinese-English speakers, that is, the same feature detectors that extract acoustic information from the speech signals spoken in their two languages, would explain that this rising F0 advantage appeared in both Mandarin and English nonwords. In contrast, for monolingual English speakers pitch differences are not crucial at the lexical level to retrieve word meanings, making irrelevant an increased sensitivity to pitch at pre-lexical stage. As a result, the top-down feedback increased the sensitivity to the acoustic saliency of rising pitch contours in the two languages of Chinese-English speakers but not when native English speakers listened to the same English nonwords, explaining the cross-language differences in the processing of acoustic salience at the pre-lexical level.

4.3 New results: Mandarin words, L1 transfer, and the bilingual lexicon

As explained in Section 4.1, results on English words from both studies showed that although native English speakers did not process pitch shape at the lexical stage, Mandarin-English speakers did. They showed a bias towards words with falling F0 shapes, which they perceived faster and more accurately, especially in prime-target pairs with identical segments. Could this falling-F0 bias of English words be explained by L1 transfer from Mandarin? Results from this experiment showed that Mandarin-English speakers perceived the falling pitch contour of T4 and rising pitch contour of T2 in Mandarin words without this falling-F0 bias (Table 1, Figure 1(b)). In fact, these speakers perceived Mandarin words with rising T2 consistently faster and less accurately than those with falling T4, revealing that the falling-F0 bias found in English words could not be explained by direct transfer from L1 Mandarin. Two questions arise. First, if L1 transfer is not a plausible explanation, then what could explain the falling-F0 bias in English words by Mandarin-English bilinguals? Second, why does the pre-lexical acoustic salience of rising contours in Chinese-English speakers' nonwords pass on to their Chinese words but not their English words?

To address these questions, we put forward the hypothesis that words are represented in phono-lexical self-organizing maps (SOMs) such as those used in BLINCS (Shook & Marian, 2013) and the Unified Model (MacWhinney, 2004), and that Chinese-English speakers differ from native English speakers in the acoustic dimensions used to store, organize, and retrieve words in their SOMs. Native English speakers obtained no significant Pitch Shape results, revealing that after an initial sensitivity to pitch differences at the pre-lexical level, pitch differences did not play a primordial role in the organization of their phono-lexical SOMs. At the lexical level, native English speakers retrieved English words based on computations of segmental similarity. In contrast, Chinese-English speakers organized words in their two languages according to both segmental and F0 acoustic similarity as showed by the significant factors of Priming Condition and Pitch Shape in word identification. Answers to the above two questions relate to how Chinese-English speakers organize words in their phono-lexical SOM according to these similarity criteria.

The first question—namely, how can the falling pitch bias in the English words of Chinese-English speakers be explained if L1 transfer from Mandarin is not a possible answer—is based on a language-specific weight hypothesis. Although Mandarin-English speakers organize Mandarin and English words in their bilingual SOM according to both segmental and pitch shape criteria, the weight attributed to, for instance, each pitch shape can be language specific. These weights regulate activation and word retrieval so that higher weights promote more activation and faster word access. In English, words with falling pitch contours are retrieved faster than words with rising pitch because they are assigned stronger weights. As explained in Section 4.1, monosyllabic and disyllabic content words in citation form have a falling pitch contour, constituting a better word representation. Moreover, previous research showed that monolingual Mandarin and Cantonese speakers show a segment preference over tone in lexical access (Sereno & Lee, 2015; Li, Lin, Wang, & Jiang, 2013). Although both information types, segment and F0, constrain lexical access as soon as segmental and F0 information becomes available in the speech signal (Malins & Joanisse, 2012), segments have a stronger effect (e.g., Sereno & Lee, 2015). Some researchers propose to differentiate this preference by attributing higher weights to segment information (e.g., Li, Lin, Wang, & Jiang, 2013). This preference would explain why the falling-F0 bias is larger at priming conditions where prime-target pairs have the same segments. Words with identical segments plus a falling pitch quickly become highly activated obtaining the fastest reaction times. In contrast, sorting computations of segmental differences between primes and targets takes time and hinders tone priming effects.

These language specific weights to particular pitch shapes also answer the second question, namely, how can we explain that this rising F0 advantage takes place in the Mandarin words but

not in the English words of Chinese-English speakers when this rising F0 advantage is passed from pre-lexical stage into the bilingual phono-lexical SOM. In Mandarin, different tones have similar weights because each tone confers lexical meaning and constitutes an equally good word representation. These similar weights among tones allow maintaining the rising F0 advantage from pre-lexical processing into the lexicon. In fact, previous research found the better representation of rising pitch in FFRs appeared in both tonal sweeps and real Mandarin words (Krishnan et al., 2001, 2004). In contrast, adding the pre-lexical rising F0 advantage to the falling-F0 bias of English words results in overriding the rising F0 advantage at lexical processing.

4.4 General discussion: The effects of tonal experience on the perception of pitch in non-tonal languages

The above results together with past research (Shook & Marian, 2016; Wang, Wang, & Malins, 2017; Ortega-Llebaria et al., 2017) provide consistent evidence supporting that experience with a tonal language models the perception of pitch in a non-tonal language. Crucially, the present experiment provided more detail about this modelling effect, revealing that it took place both at pre-lexical processing via feedback from the lexicon, and in the lexicon itself.

Contrary to the widespread assumption that the F0 variation of the speech signal is not relevant for spoken word recognition in non-tonal languages such as English, our results showed that for both native English speakers and tonal speakers of English, the F0 variation present in English words and nonwords had a processing cost revealing that F0 variation was indeed processed at the pre-lexical level across languages and speakers. However, Chinese-English speakers, but not native English speakers, showed an enhanced sensitivity to the acoustic salience of rising pitch at the pre-lexical level. The important role of pitch in the lexicon of tonal speakers feeds back to low-level acoustic processing, enhancing their pre-lexical sensitivity to pitch in both Mandarin and English targets. Although a general increased sensitivity to pitch by tonal speakers in comparison to non-tonal speakers has been well documented (e.g., Bidelman, Hutka, & Moreno, 2013; Krishnan, Gandour, & Suresh, 2014), our contribution is to show that this increased sensitivity to pitch affects their non-tonal language at the pre-lexical level because of the effect of tonal experience on pre-lexical processing.

After pre-lexical processing, only Chinese-English speakers continue processing pitch at the lexical level in both their tonal and non-tonal languages. The falling-F0 bias on English words by Chinese-English speakers but not by non-tonal speakers obtained in both experiments constitutes strong supporting evidence that Chinese-English speakers encoded English words with a falling pitch contour in their lexicon. This falling pitch contour can be related with the consistently reported production patterns of Chinese-English speakers summarized in the Introduction. Recall that Mandarin and Cantonese speakers of English tended to produce stressed syllables in English and Spanish words with an H* pitch accent, which remained unchanged despite the requirements of sentence intonation. For example, in Yes-No questions the last stressed syllable takes on a L* pitch accent followed with an H% boundary tone, such as Ma- in “Mary?” is realized with an L* and -ry with an H%. Chinese speakers tended to produce Ma- with an H* and then squeeze L*H% on -ry (e.g., McGory, 1997). Similarly, the deaccentuation required in the reporting sentences and post-focal clauses of English and Spanish was not fully realized by Chinese speakers, causing a lack of contrast between the word in focus and the non-focused clause, which triggered several repair strategies. For example, Chinese speakers of English tended to insert a pause after the focal word to differentiate it from the following post-focal clause, and to overshoot the H* in the focal word (Ortega-Llebaria & Colantoni, 2014). As for alignment, this H* remains in the stressed syllables of

Spanish words in declarative sentences despite the fact that Spanish speakers produce in this context a post-tonic peak described as an L*H pitch accent (Chen, 2007). Altogether, these production patterns show that Chinese speakers of English (and Spanish) consistently preserve an H* on the stressed syllable of English words.

The consistent production of this H* peak in the stressed syllables of the English words despite sentence intonation requirements together with the consistent falling-F0 bias showing that Chinese-English speakers represented English words with a falling pitch in their lexicon suggest that for Chinese speakers, English works as a one-tone language. Because there is only one tone—namely, H* aligned with the stressed syllable—tonal shape does not change meanings. However, as in any tonal language, H* will be preserved despite sentence intonation requirements. Although it can be expanded in range as long as it preserves its H* shape as it has been documented in the H* overshoot of English words in focus (e.g., Ortega-Llebaria & Colantoni, 2014), any sentence intonation requirement that modifies the H* contour—namely, the deaccentuation of reporting sentences, the L* in Yes-No Questions, or the delayed peak alignment L*H in Spanish declarative sentences—will be banned. As a result, English sentence intonation is re-interpreted to preserve this H* tone, for example, the Yes-No Question tune L*H% is re-interpreted as H* L-H%. For Chinese-English speakers to process intonation as native English speakers do, this tonal re-interpretation of English intonation needs to be unlearned. The falling-F0 bias has to disappear and the effect of pitch shape in English words has to become statistically non-significant. Whether and how this goal can be achieved is a question for future research.

5 Conclusion

Previous research showed the tonal experience of Chinese-English speakers modulated the perception of pitch in the words of their non-tonal language. The present study expanded these results by showing the tonal experience of Chinese-English speakers modulated the perception of pitch in their non-tonal language at both pre-lexical and lexical processing. At a pre-lexical stage, both Chinese-English speakers and native English speakers processed pitch differences. However, only Chinese-English speakers showed an increased sensitivity to the acoustic salience of rising in their two languages. This increased sensitivity was due to their tonal experience: the important role of pitch in the lexicon of tonal speakers feeds back to low-level acoustic processing, enhancing their pre-lexical sensitivity to pitch via shared feature detectors. As a result, this increased sensitivity due to their tonal experience affected not only their tonal but also their non-tonal language.

Chinese-English speakers, but not native English speakers, continued processing pitch at the lexical level, and they did it in their two languages. The falling-F0 bias in English words consistently obtained in this experiment and by Ortega-Llebaria et al. (2017) provided strong evidence supporting that Chinese-English speakers encoded pitch in English words and that a falling pitch contour constituted a preferred word representation. This bias in perception was related to the consistent production of an H* peak in the stressed syllable of English words. Chinese-English speakers tend to maintain that peak regardless of the requirements of sentence intonation. Altogether, this cumulative evidence suggests the tonal experience of these speakers makes them re-interpret English as a one-tone language, where the only tone, H*, is preserved in every content word, and as a result, English intonation is re-interpreted in function of this tone preservation.

Funding

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

ORCID iD

Marta Ortega-Llebaria  <https://orcid.org/0000-0003-1637-3541>

Supplemental material

Supplemental material for this article is available online.

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 48.
- Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, 39, 445–459.
- Beckman, M. E. (1986). *Stress and non-stress accent* (Vol. 7). Dordrecht-Holland/Riverton, USA: Foris Publications.
- Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PloS one*, 8, e60676.
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer program]. Version 6.0.39, retrieved from <http://www.praat.org/>
- Burnham, D., Singh, L., Mattock, K., Woo, P. J., & Kalashnikova, M. (2017). Constraints on tone sensitivity in novel word learning by monolingual and bilingual infants: Tone properties are more influential than tone familiarity. *Frontiers in Psychology*, 8, 1–14.
- Calhoun, S., Wollum, E., & Kruse, E. (this volume). Prosodic prominence and focus: Expectation affects interpretation in Samoan and English. *Language and Speech. Special Issue on Prosody*. (Ed. Martine Grice). Advance online publication. <https://journals.sagepub.com/doi/full/10.1177/0023830919890362>
- Chen, Y. (2007). *A comparison of Spanish produced by Mandarin L2 learners and native speakers—an acoustic phonetics approach*. University of Illinois at Urbana-Champaign.
- Coffey, E. B., Herholz, S. C., Chepesiuk, A. M., Baillet, S., & Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature communications*, 7, 11070.
- Cutler, A., & Norris, D. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology*. 1988, 14, 113–121.
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language & Cognition*, 1, 131–149.
- Gussenhoven, C. (2014). On the intonation of tonal varieties of English. *The Oxford Handbook of World Englishes online*.
- Gussenhoven, C., & Udofot, I. (2010). Word melodies vs. pitch accents: A perceptual evaluation of terracing contours in British and Nigerian English.
- Gut, U. B. (2008). Nigerian English: Phonology. *Varieties of English*, 4, 35–54.
- Intartaglia, B., White-Schwoch, T., Meunier, C., Roman, S., Kraus, N., & Schön, D. (2016). Native language shapes automatic neural processing of speech. *Neuropsychologia*, 89, 57–65.
- Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Preference for the predominant stress patterns of English words. *Child Development*, 64, 675–687.
- Kember, H., Choi, J., Yu, J., & Cutler, A. (this volume). The processing of linguistic prominence. *Language and Speech. Special Issue on Prosody*. (Ed. Martine Grice). Advance online publication. <https://journals.sagepub.com/doi/full/10.1177/0023830919880217>
- Krishnan, A., & Parkinson, J. (2000). Human frequency-following response: Representation of tonal sweeps. *Audiology and Neurotology*, 5, 312–321.
- Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequency-following response: Representation of pitch contours in Mandarin tones. *Hearing Research*, 189, 1–12.
- Krishnan, A., Gandour, J. T., & Suresh, C. H. (2014). Cortical pitch response components show differential sensitivity to native and nonnative pitch contours. *Brain and Language*, 138, 51–60.

- Krishnan, A., Gandour, J. T., Ananthakrishnan, S., & Vijayaraghavan, V. (2014). Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. *Neuropsychologia*, 59, 1–12.
- Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing: Looking back and to the future. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 531–553). New York: Oxford University Press.
- Lee, C. Y. (2007). Does horse activate mother? Processing lexical tone in form priming. *Language and Speech*, 50, 101–123.
- Lee, L., & Nusbaum, H.C. (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin. *Perception & Psychophysics*, 53, 157–165.
- Lewy, N. (2008). The Lewy and Grosjean BIMOLA model. In F. Grosjean (Ed.), *Studying Bilinguals* (pp. 201–210). New York: Oxford University Press.
- Li, C., Lin, C. Y., Wang, M., & Jiang, N. (2013). The activation of segmental and tonal information in visual word recognition. *Psychonomic Bulletin & Review*, 20, 773–779.
- Luchkina, T., & Cole, J. (this volume). Perception of word-level prominence in free word order language discourse. *Language and Speech. Special Issue on Prosody*. (Ed. Martine Grice). Advance online publication. <https://journals.sagepub.com/doi/full/10.1177/0023830919884089>
- McGory, J. T. (1997). *Acquisition of intonational prominence in English by Seoul Korean and Mandarin speakers* (Doctoral dissertation) The Ohio State University, Ohio, USA.
- MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition* (pp. 249–308). New York: Lawrence Erlbaum Associates, Inc.
- MacWhinney, B. (2005). A unified model of language acquisition. *Handbook of Bilingualism: Psycholinguistic Approaches*, 49–67.
- 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, 2032–2043.
- Miller, J. L. (1978). Interactions in processing segmental and suprasegmental features of speech. *Perception & Psychophysics*, 24, 175–180.
- Neergaard, K. D., Xu, H., & Huang, C. (2016). Database of Mandarin Neighborhood Statistics. *Proceedings of the Language Resources and Evaluation Conference (LREC 2016)*, Portoroz, Slovenia.
- Ortega-Llebaria, M., & Colantoni, L. (2014). L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in second language acquisition*, 36, 331–353.
- Ortega-Llebaria, M., Nemogá, M., & Presson, N. (2017). Long-term experience with a tonal language shapes the perception of intonation in English words: How Mandarin–English bilinguals perceive “Rose?” vs. “Rose”. *Bilingualism: Language and Cognition*, 20, 367–383.
- Pierrehumbert, J. B. (1980). The phonology and phonetics of English intonation. (Dissertation) Massachusetts Institute of Technology, Massachusetts, USA.
- R Core Team (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna*. Retrieved from URL <http://www.R-project.org/>.
- Repp, B. H., & Lin, H. B. (1990). Integration of segmental and tonal information in speech perception: A cross-linguistic study. *The Journal of the Acoustical Society of America*, 87, S46–46.
- Sereno, J. A., & Lee, H. (2015). The contribution of segmental and tonal information in Mandarin spoken word processing. *Language and Speech*, 58, 131–151.
- Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language & Cognition*, 16, 304–324.
- Shook, A., & Marian, V. (2016). The influence of native-language tones on lexical access in the second language. *The Journal of the Acoustical Society of America*, 139, 3102–3109.
- Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, 124, 128–142.
- Todaka, Y. (1990). Error Analysis of Japanese Students’ Intonation and its Prosodic Analysis. (Master’s thesis) University of California, Los Angeles, USA.

- Tseng, C. Y., Su, C. Y., & Visceglia, T. (2013). Under-differentiation of English lexical stress contrasts by L2 Taiwan speakers. In *Speech and Language Technology in Education Conference*, Wisconsin Dells, WI, 9–11 December 2013.
- Tremblay, A. (2008). Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics*, 29, 553–584.
- Visceglia, T., & Fodor, J. D. (2006). Fundamental frequency in Mandarin and English. *Interfaces in multilingualism: Acquisition and representation*, 4, 27–59.
- Visceglia, T., Su, C. Y., & Tseng, C. Y. (2012). Comparison of English narrow focus production by L1 English, Beijing and Taiwan Mandarin speakers. In *Speech Database and Assessments (Oriental COCODA), 2012 International Conference*, Kerala, India (pp. 47–51). *IEEE*.
- Wang, X., Wang, J., & Malins, J. G. (2017). Do you hear ‘feather’ when listening to ‘rain’? Lexical tone activation during unconscious translation: Evidence from Mandarin-English bilinguals. *Cognition*, 169, 15–24.
- Wayland, R., Zhu, Y. Q., & Kaan, E. (2015). Perception of pitch contours by native and non-native tone listeners. In *Proceedings of the 18th International Congress of Phonetic Sciences, ICPhS 2015*, Glasgow, UK, 10–14 August 2015. University of Glasgow.