Original Article

Language and Speech

The Influence of Tonal and Atonal Bilingualism on Children's Lexical and Non-Lexical Tone Perception

Language *and* Speech
1–21
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0023830919834679
journals.sagepub.com/home/las
SAGE

Laura M. Morett
University of Alabama, USA

Abstract

This study examined how bilingualism in an atonal language, in addition to a tonal language, influences lexical and non-lexical tone perception and word learning during childhood. Forty children aged 5;3–7;2, bilingual either in English and Mandarin or English and another atonal language, were tested on Mandarin lexical tone discrimination, level-pitch sine-wave tone discrimination, and learning of novel words differing minimally in Mandarin lexical tone. Mandarin–English bilingual children discriminated between and learned novel words differing minimally in Mandarin lexical tone more accurately than their atonal–English bilingual peers. However, Mandarin–English and atonal–English bilingual children discriminated between level-pitch sine-wave tones with similar accuracy. Moreover, atonal–English bilingual children showed a tendency to perceive differing Mandarin lexical and level-pitch sine-wave tones as identical, whereas their Mandarin–English peers showed no such tendency. These results indicate that bilingualism in a tonal language in addition to an atonal language—but not bilingualism in two atonal languages—allows for continued sensitivity to lexical tone beyond infancy. Moreover, they suggest that although tonal–atonal bilingualism does not enhance sensitivity to differences in pitch between sine-wave tones beyond infancy any more effectively than atonal–atonal bilingualism, it protects against the development of biases to perceive differing lexical and non-lexical tones as identical. Together, the results indicate that, beyond infancy, tonal–atonal bilinguals process lexical tones using different cognitive mechanisms than atonal–atonal bilinguals, but that both groups process level-pitch non-lexical tone using the same cognitive mechanisms.

Keywords
Cross-linguistic speech perception, child bilingualism, lexical tone, pitch processing, word learning

1 Introduction

A majority of people in the world speak tonal languages, which use pitch to differentiate between lexical semantics or morphosyntax (Fromkin, 1978)—a property known as lexical tone. Particularly

Corresponding author:
Laura M. Morett, University of Alabama, Box 870240, Tuscaloosa, AL 35401, USA.
Email: lmorett@ua.edu

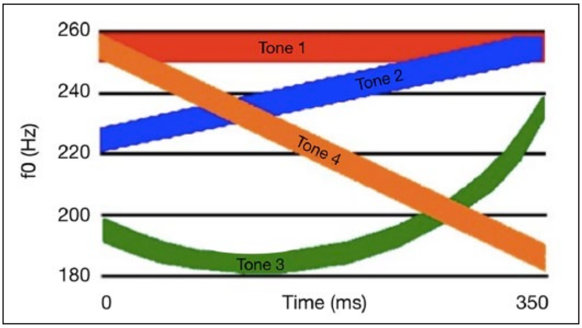

Figure 1. Stylized pitch contours of the four principal Mandarin lexical tones.

among children, a majority of the speakers of tonal languages are bilingual, and a majority of these bilinguals speak a second language that is atonal (Feng, 2007). Despite growing interest in how experience with tonal languages, in addition to atonal languages, affects perception of lexical tone, at present, little is known about how tonal–atonal bilingualism affects non-lexical tone perception, particularly in childhood. Moreover, despite evidence that bilingualism in two atonal languages extends the developmental duration of sensitivity to lexical tone (Graf Estes & Hay, 2015), no published research to date has examined whether its effect on lexical tone perception is similar to that of tonal–atonal bilingualism. Finally, relatively little is known about how perception of both lexical and non-lexical tone develops beyond infancy. The goal of the current research was to fill these lacunae by investigating the effect of bilingualism in an atonal language in addition to a tonal language, as well as bilingualism in two atonal languages, on lexical and non-lexical tone perception in early childhood. As a result, this work provides insight into how tonal–atonal and atonal–atonal bilingualism influence the development of lexical and non-lexical tone perception beyond infancy.

1.1 Lexical tone perception and development

In order to understand the development and mature state of lexical tone perception, it is important to understand the acoustic properties of lexical tone in general, as well as the specifics of Mandarin lexical tone in particular. Unlike segmental phonemes such as consonants and vowels, lexical tone is a suprasegmental, consisting of a pitch superimposed upon a syllable or word (Xu, 1999). Lexical tone is primarily identified by fundamental frequency (F0) cues, including mean pitch and pitch contour (Gandour, 1983, 1984; Massaro, Cohen, & Tseng, 1985), and is secondarily identified by cues in duration, amplitude, vocal range, and register (Blicher, Diehl, & Cohen, 1990; Liu & Samuel, 2004). Thus, lexical tone is perceived distinctly from segmental phonology and is classified based on these cues according to a fixed set of categories in a given tonal language (Francis, Ciocca, & Ng, 2003; Xi, Zhang, Shu, Zhang, & Li, 2010). Mandarin, the tonal language of interest in the current study, is a contour tone system, meaning that its tones are distinguished from one another by the pitch trajectory of their F0 (Chao, 1965). There are four principal categories of lexical tones in Mandarin: high flat (1), rising (2), low dipping (3), and falling (4) (see Figure 1).[1] In Mandarin, lexical tone differentiates lexically rather than inflectionally. For native Mandarin speakers, lexical tone constrains activation of word referents in parallel with segmental phonology (Malins & Joanisse, 2010, 2012), indicating that lexical tone serves as a critical cue to meaning in Mandarin. The current study focuses on Mandarin because of the high information load of its lexical tones (Oh, Pellegrino, Coupé, & Marsico, 2013; Surendran & Levow, 2004), which may influence how overlapping non-lexical tones are perceived.

Research on the neural bases of lexical tone perception has provided insight into how language experience affects the mechanisms used to process lexical tone. In contrast to tonal language speakers, who process lexical tone in their native language using mainly left-lateralized brain regions involved in speech perception, speakers of atonal languages such as English and other tonal languages process lexical tone using right-lateralized brain regions as well (Gandour, Wong, & Hutchins, 1998; Hsieh, Gandour, Wong, & Hutchins, 2001; Klein, Zatorre, Milner, & Zhao, 2001; Wang, Jongman, & Sereno, 2001; Xu, 1999), which are also used to process affective intonation (Tompkins & Flowers, 1985; Wildgruber, Pihan, Ackermann, Erb, & Grodd, 2002) as well as non-lexical tone (Tervaniemi et al., 1999, 2000). When a tonal language is acquired, the laterality of processing lexical tones from that language shifts to the left hemisphere (Wang, Sereno, Jongman, & Hirsch, 2003; Wang, Spence, Jongman, & Sereno, 1999), predicting lexical tone acquisition efficacy (Wong, Perrachione, & Parrish, 2007). Taken together, these findings indicate that, in a familiar tonal language, lexical tone processing relies upon neural mechanisms used to process familiar segmental phonemes, whereas in an unfamiliar tonal language, lexical tone processing relies upon neural mechanisms used to process affective prosody and non-lexical tone.

In light of evidence suggesting that tonal and atonal language speakers rely on different neural mechanisms to process lexical tone, it follows that lexical tone perception should develop differently in children with experience exclusively with tonal and atonal languages. Indeed, there is evidence of developmental divergence in lexical tone perception based on language experience, and that the trajectory of lexical tone development parallels, but is more complex—and protracted—than segmental phonological development (Singh & Fu, 2016). While all children are born with the ability to discriminate between lexical tones in any language, similar to segmental phonemes, perceptual reorganization takes place between the age of 0;4 and 0;9, such that lexical tone discrimination begins to improve in children exposed exclusively to a tonal language and to decline in children exposed exclusively to an atonal language (Harrison, 2000; Liu & Kager, 2014; Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; Tsao, 2017; Yeung, Chen, & Werker, 2013). Children with exclusive experience with an atonal language continue to demonstrate some elasticity in lexical tone perception until age 2;0 (Hay, Graf Estes, Wang, & Saffran, 2015; Liu & Kager, 2014), after which they fail to differentiate between lexical tones (Quam & Swingley, 2010; Singh, Hui, Chan, & Golinkoff, 2014).

Unlike children with experience exclusively with an atonal language, children with experience exclusively with a tonal language demonstrate protracted development of lexical tone perception far beyond infancy. Similar to children with experience exclusively with an atonal language, children with experience exclusively with a tonal language discriminate with a high rate of accuracy between words differing minimally in lexical tone in early childhood (Singh, Poh, & Fu, 2016; Wong, Schwartz, & Jenkins, 2005; Yuet Sheung Lee, Chiu, & van Hasselt, 2002). However, sensitivity to familiar lexical tones continues to improve during childhood in children with experience exclusively with a tonal language (Lee et al., 2012; Singh, Tan, & Wewalaarachchi, 2017). Indeed, there is evidence that discrimination of familiar lexical tones does not reach adult accuracy until 10;0 (Ciocca & Lui, 2003; Wong & Leung, 2018).

Given that children with experience exclusively with tonal and atonal languages exhibit divergent developmental trajectories for lexical tone perception, the question arises of how lexical tone perception develops in children bilingual in both a tonal and an atonal language. Interest in this question is growing, and while the results of several recent studies are somewhat mixed, they nevertheless suggest that tonal–atonal bilinguals exhibit prolonged flexibility in lexical tone perception during childhood. It is not until approximately 1;0 that tonal–atonal bilinguals interpret suprasegmental pitch correctly as a phonemic contrast in Mandarin and a non-phonemic acoustic attribute in English (Singh & Foong, 2012; Singh et al., 2016). During early childhood, sensitivity to lexical tone appears to follow a complex, protracted trajectory in tonal–atonal bilinguals, such that it decreases early on (Wewalaarachchi, Wong, & Singh, 2017) but increases and becomes more nuanced with age (Singh, Goh, & Wewalaarachchi, 2015). It is important to note, however, that children with experience in two atonal languages retain sensitivity to lexical tone (Graf Estes & Hay, 2015) as well as unfamiliar segmental phonemes (Singh, 2018) for longer than children with experience in only one atonal language. This suggests that the protracted sensitivity to lexical tone observed in tonal–atonal bilinguals may result from bilingualism rather than from experience with an atonal language, per se. Nevertheless, lexical tone perception in children with experience with both a tonal and an atonal language indicates that it reflects both similarities and differences relative to that of children with experience exclusively with a single tonal language.

1.2 Non-lexical tone perception and development

Unlike lexical tones, which are constrained according to the conventions of the language in which they occur, non-lexical tones range from those with level pitches to those with pitch contours. In general, research indicates that tonal language speakers perceive non-lexical tones with pitch contours consistently and accurately when they are similar to the contours of the lexical tones in their language. More specifically, tonal language speakers show evidence of categorical perception of non-lexical tones with pitch contours, with decreased sensitivity to those falling within the acoustic boundaries corresponding to lexical tones in their language relative to those straddling these categories, whereas atonal language speakers show evidence of psychophysical perception of all tones with pitch contours (Bent, Bradlow, & Wright, 2006; Chandrasekaran, Krishnan, & Gandour, 2007, 2009; Hallé, Chang, & Best, 2004; Lee, Vakoch, & Wurm, 1996; Peng et al., 2010; Xu, Gandour, & Francis, 2006). Findings concerning the influence of tonal language experience on discrimination of non-lexical tones with level pitches is mixed, with some work showing evidence of perceptual enhancement (Giuliano, Pfordresher, Stanley, Narayana, & Wicha, 2011; Pfordresher & Brown, 2009), some work showing evidence of perceptual suppression (Stagray & Downs, 1993; Tanner & Rivette, 1964), and some work showing no effect (Bent et al., 2006; Burns & Sampat, 1980). Together, these findings suggest that tonal language experience influences non-lexical tone perception and that its influence may depend upon on the similarity of the pitches of non-lexical tones to the pitches of familiar lexical tones.

In addition to the research discussed above, a number of studies provide evidence of a bi-directional relationship between experience and perception of musical pitch and lexical tone. Work examining the effect of tonal language experience on musical pitch perception indicates that fluency in tonal languages is positively associated with absolute pitch prevalence (Deutsch, Dooley, Henthorn, & Head, 2009; Deutsch, Henthorn, Marvin, & Xu, 2006) as well as enhanced musical pitch perception (Wong et al., 2012). Moreover, tonal language speakers with amusia show impaired discrimination and identification of non-lexical pitch contours (Jiang, Hamm, Lim, Kirk, & Yang, 2010; Liu et al., 2012). Work examining the effect of musical experience on lexical tone perception indicates that speakers of atonal languages with significant musical experience learn lexical tone more effectively than speakers of atonal languages without significant musical experience (Delogu, Lampis, & Belardinelli, 2006, 2010; Lee & Hung, 2008). Furthermore, atonal language speakers with amusia show impairments in their ability to learn lexical tone (Tillmann et al., 2011), and tonal language speakers with amusia show impairments in lexical tone discrimination and identification (Nan, Sun, & Peretz, 2010). More recent work integrating these two lines of inquiry suggests that tonal language and musical experience exert independent effects on musical pitch perception (Bidelman, Gandour, & Krishnan, 2011a, 2011b; Bidelman, Hutka, & Moreno, 2013), with musical experience exerting a larger effect than tonal language experience (Hutka, Bidelman, & Moreno, 2015).

Despite the extensiveness of extant research examining the influence of tonal and atonal language experience on non-lexical tone processing in adults, relatively little research has examined it in children. Moreover, the findings of this research are mixed, suggesting both parallels and differences between the development of lexical and non-lexical tone perception. In infancy, exposure to both English and Mandarin permits successful discrimination between non-lexical tones with pitch contour analogues of Thai lexical tones, whereas exposure only to English results in a developmental decline in sensitivity to the differences between these tones (Mattock & Burnham, 2006). By contrast, exposure to Dutch, an atonal language, hinders discrimination between Mandarin lexical tones and analogous three-note non-lexical tones at 0;4 but not 1;0 (Chen & Kager, 2016). In early childhood, 3- to 5-year-old Mandarin-speaking children discriminate between non-lexical tones with rising and falling pitch contours more accurately than their English-speaking peers (Creel, Weng, Fu, Heyman, & Lee, 2018). However, English-speaking children aged 5;0, 6;0, and 8;0 and 18+ are better able to discriminate between non-lexical tones with level pitches and pitch contour analogues of Thai lexical tones than Thai lexical tones (Burnham, Francis, & Webster, 1996). Together, these findings provide evidence that experience with tonal languages may enhance the development of non-lexical tone perception.

Given that work examining non-lexical tone development beyond infancy has only tested children with experience with either a single tonal or atonal language (Burnham et al., 1996; Creel et al., 2018), it is unclear how non-lexical tone is perceived by children with experience with two languages beyond infancy. One possibility is that experience with two languages does not affect non-lexical tone perception beyond infancy. This possibility is consistent with research indicating that monolingual tonal and atonal language speakers discriminate between non-lexical tones with level pitches with similar accuracy (Bent et al., 2006; Burns & Sampat, 1980), despite showing differences in lexical tone discrimination accuracy. On the surface, this possibility appears consistent with theories of speech perception that hold that speech and non-speech sounds are processed and represented differently (Liberman & Mattingly, 1989). However, findings indicating that tonal language speakers perceive non-lexical tones with level pitches, non-lexical tones with pitch contours, and lexical tones along a speech-like-ness spectrum (e.g., Bent et al., 2006) suggest that lexical and non-lexical tone processing may not be strictly modular or dichotomous.

An alternative possibility is that experience with two languages affects non-lexical tone perception beyond infancy and that its effects depend on language tonality. If this is the case, tonal–atonal bilinguals should differ from atonal–atonal bilinguals in their perception of distinctions between non-lexical tones. This possibility is consistent with research demonstrating that tonal monolinguals discriminate more accurately than atonal monolinguals between level-pitched non-lexical tones with small inter-pitch intervals (Giuliano et al., 2011) or that straddle the F0 boundaries of lexical tones from their native language (Stagray & Downs, 1993; Tanner & Rivette, 1964). Furthermore, it is consistent with research suggesting that tonal monolinguals are more likely than atonal monolinguals to possess absolute pitch (Deutsch et al., 2006, 2009) and to perceive musical pitch more accurately (Wong et al., 2012), as well as to show adverse effects of amusia on lexical tone perception (Jiang et al., 2010; Liu et al., 2012). Theoretically, this possibility is consistent with models of speech perception that hold that speech and non-speech sounds are processed and represented similarly using shared mechanisms (Kluender, Diehl, & Killeen, 1987; Kuhl, 1981). Specifically, the finding that tonal–atonal bilinguals perceive non-lexical tone differently than atonal–atonal bilinguals beyond infancy would suggest that speaking a tonal language in addition to an atonal language influences non-lexical tone perception, demonstrating transfer from production to perception as well as from lexical to non-lexical tone perception similar to transfer observed in tonal monolinguals. Moreover, it would demonstrate that bilingualism in a tonal language in addition to an atonal language does not completely block this transfer beyond infancy.

1.3 The current study

The primary objective of the present study was to investigate how experience with two languages varying in tonality affects lexical and non-lexical tone perception beyond infancy. To achieve this objective, the present study examined perception of lexical and non-lexical tone by young children bilingual in Mandarin and English (tonal–atonal bilinguals) and bilingual in English and another atonal language (atonal–atonal bilinguals). These two groups were compared so that the influences of experience with two languages, regardless of tonality, and experience with a tonal language could be disentangled with respect to lexical and non-lexical tone perception. It is important to disentangle the influences of bilingualism and tonal language experience on tone perception because the majority of children who grow up speaking Mandarin learn to speak English in early childhood (Feng, 2007). Thus, theories of lexical and non-lexical tone perception and development must account for the influences of both bilingualism in general and tonal language experience in particular to ensure that they are ecologically valid.

With respect to lexical tone perception, it was predicted that tonal–atonal bilinguals would discriminate more accurately than atonal–atonal bilinguals. This prediction was based on evidence that, beyond infancy, tonal monolinguals' ability to discriminate between familiar lexical tones improves (Ciocca & Lui, 2003; Lee et al., 2012), whereas atonal monolinguals' lexical tone discrimination ability declines (Quam & Swingley, 2010; Singh et al., 2014). Although atonal–atonal bilinguals retain the ability to discriminate between lexical tones for longer than atonal monolinguals, atonal–atonal bilinguals' lexical tone discrimination declines to the level of atonal monolinguals by 2;2, indicating that atonal–atonal bilinguals' lexical tone discrimination does not persist beyond infancy (Graf Estes & Hay, 2015). While this prediction is at odds with work demonstrating that tonal–atonal bilinguals show reduced sensitivity beyond infancy to lexical tone substitutions in words relative to their earlier sensitivity (Singh et al., 2015; Wewalaarachchi et al., 2017), it is consistent with findings from tasks requiring explicit discrimination between lexical tones (Bent et al., 2006; Hallé et al., 2004; Lee et al., 1996), which were used in the present study.

With respect to level-pitched non-lexical tone perception, it was predicted that tonal–atonal bilinguals would discriminate with accuracy similar to that of atonal–atonal bilinguals. This prediction was based on evidence that, beyond infancy, non-lexical tone discrimination develops similarly in tonal and tonal monolinguals (Burnham et al., 1996). This prediction was also consistent with the findings of work examining perception of level-pitched non-lexical tones, which was tested in the present study, indicating that tonal and atonal bilinguals discriminate between these tones with similar accuracy (Bent et al., 2006; Burns & Sampat, 1980). Given that no published research to date has examined non-lexical tone perception in atonal–atonal bilinguals, the present study provides insight into whether superior phonemic perception in bilinguals arises from—or affects—acoustic perception, thereby informing theories of bilingual speech perception.

2 Method

2.1 Participants

Forty children aged 5;3–7;2 (20 females, 20 males; average age = 72.0 months, *SD* = 5.0 months) were recruited from preschools across Singapore with the permission of their caregivers to participate in this study. All children were bilingual in English and Mandarin (*n* = 19) or

Table 1. Age and L2 experience and proficiency of participants by language group.

Attribute	Language group			
	Atonal–English	Mandarin–English	*t*	*p*
Age (months)	71.80 (4.89)	71.88 (5.24)	–0.04	0.97
Percentage of time L2 spoken by main caregiver	69.00 (24.09)	50.33 (44.06)	1.48	0.15
Length of time of L2 experience (months)	69.00 (7.30)	67.47 (13.18)	0.40	0.69
L2 comprehension rating (1–7)	5.45 (1.28)	5.25 (1.29)	0.46	0.65
L2 production rating (1–7)	5.45 (1.64)	4.88 (1.59)	1.07	0.29

English and a second atonal language (Malay, Tamil, Hindi, Punjabi, or Telegu; $n = 21$). In all cases, English was learned first and was dominant relative to the other language. Most children had experience with the language other than English from birth; three children from each group gained experience with the non-English language between birth and 3;0.[2] Table 1 provides further information concerning children's language experience based on a caregiver-report language background questionnaire. In this questionnaire, primary caregivers of children reported their child's age (in months), estimated the percentage of time that they spoke each language in the presence of their child, and reported the duration (in months) of their child's experience with each language. Additionally, caregivers rated their child's comprehension and production of each language on a 1–7 scale (1 = poor; 7 = very good). Responses to this questionnaire revealed that tonal–atonal and atonal–atonal bilinguals did not differ significantly with respect to age, duration of language experience, or L2 proficiency. The questionnaire did not include any questions concerning children's musical experience. No speech or hearing problems were reported by any of the children or their caregivers, and all children had normal or corrected-to-normal vision.

2.2 Stimuli and procedure

Prior to experimental sessions, participants' caregivers completed the language background questionnaire. Data from the language background questionnaire were used to classify participants according to whether their language other than English was tonal or atonal.

In experimental sessions, participants performed three tasks: a non-lexical tone discrimination task, a lexical tone discrimination task, and a word learning task. Testing was conducted on an individual basis in a quiet room in the children's preschool. In all tasks, audio stimuli were played diotically over headphones at a volume judged to be comfortable by the experimenter (the author). Participants completed a brief practice session with feedback before beginning each task, but no feedback was provided during experimental trials. The order of the non-lexical and lexical tone discrimination tasks was counterbalanced between participants, but the word learning task was always presented last due to its difficulty. During experimental trials, all responses were entered into the experimental control program by the experimenter using a coding scheme devised prior to the experiment. Due to the presentation of auditory stimuli to participants through headphones, the experimenter was unaware of which stimuli were being presented.

2.2.1 Non-lexical tone discrimination task. Non-lexical tone stimuli consisted of 15 level-pitched sine wave tones varying in frequency from 120 to 162 Hz. This pitch range covered the average F0 range for male English speakers (Hillenbrand, Getty, Clark, & Wheeler, 1995), taking into account

Table 2. Words used in lexical discrimination task.

Word	Token			
	High flat tone (1)	Rising tone (2)	Low dipping tone (3)	Falling tone (4)
bao[a]	bāo[a]	báo[a]	bǎo[a]	bào[a]
ba	bā	bá	bǎ	bà
di	dī	dí	dǐ	dì
gu	gū	gú	gǔ	gù
lei	lēi	léi	lěi	lèi
mao	māo	máo	mǎo	mào
wei	wēi	wéi	wěi	wèi

[a]Used in practice trials only.

that the average fundamental frequency for Mandarin speakers is lower than that of English speakers (Chao & Pian, 1948). As in other studies testing discrimination of level-pitched non-lexical tones (Burnham et al., 1996; Stagray & Downs, 1993), inter-tone pitch spacing of 3 Hz was used because of its congruence with perceptual thresholds.

In each experimental trial, a standard stimulus of 141 Hz was always presented first. This was followed by a 2000 ms interstimulus interval, which was followed by a test stimulus. Participants were instructed to say "same" if the two tones were identical, "different" if the tones differed, or "I don't know" if they couldn't tell. The practice session consisted of three trials (test stimuli: 1 same frequency + 1 lower frequency + 1 higher frequency), and the experimental task consisted of 28 trials (test stimuli: 14 same frequency + 14 different frequency; one trial per frequency differing from 141 Hz with an equal number of "same" trials).

2.2.2 Lexical tone discrimination task. Lexical tone stimuli consisted of six sets of four monosyllabic Mandarin words differing minimally in lexical tone, yielding a total of 24 stimuli (see Table 2). Real Mandarin words, rather than pseudo-words, were used in this task. Although the use of real Mandarin words results in a confound between lexical familiarity and lexical tone discrimination, real words were used because of concerns about possible F0 differences between pseudo-words and real Mandarin words that may have affected performance. All words consisted of consonants and vowels present in both Mandarin and English, and word sets varied minimally in lexical tone, encompassing the four principal lexical tones in Mandarin: high flat (1), rising (2), low dipping (3), or falling (4) (Chao, 1965). Prior to the experiment, a Singaporean female native speaker of Mandarin was recorded reading the words, and tokens were edited for consistency of onset/offset duration.

In this task, a discrimination paradigm was used to test lexical tone perception. Unlike tone identification paradigms, which require participants to label the lexical tone presented (e.g., 1 or high flat), atonal language speakers can perform the discrimination task successfully without training or explicit knowledge of the tone system of the target language. Thus, this task represents a more conservative test of perceptual differentiation than tone identification. In each trial, a randomly-selected token from a set of words differing minimally in lexical tone was presented. After an interstimulus interval, either the same token or another token from the same word set that differed minimally from the first token in lexical tone was presented. Participants were instructed to say "same" if the tokens were the same, and "different" if the tokens differed. The practice session for this task consisted of four trials (2 same pairs + 2 different pairs from a word set not presented

Figure 2. Images used to represent referents of the six novel words.

in experimental trials), and the experimental task consisted of 36 trials (18 same pairs + 18 different pairs).

2.2.3 Word learning task. Auditory stimuli consisted of two triads of three novel words formulated according to the rules of English phonology (*/nuθ/* and */jis/*). Words in each triad varied minimally in lexical tone, which was either high flat (1), rising (2), or falling (4).[3] Another female Singaporean Mandarin–English bilingual speaker was recorded reading the words, and recordings were edited for consistency of onset/offset duration. Referents used in this task consisted of six color images of novel objects (see Figure 2). Correspondences between visual referent representations and words were counterbalanced between participants.

This task consisted of two blocks presented in counterbalanced order, each of which comprised one triad. Each block consisted of three phases, presented in the following order: learning, practice, and test. In learning trials (three counterbalanced repetitions; nine trials), words in each triad were presented sequentially, with each word presented 1000 ms after the onset of its referent, which remained on the screen for an additional 4000 ms, encompassing word presentation (total duration: 5000 ms). In practice trials (one per word; three trials), participants heard each word in the triad sequentially, during which two referents initially presented in learning trials were displayed on either side of the screen. Participants indicated the referent corresponding to each word by pointing at it and received feedback indicating whether their response was correct or incorrect. Test trials (two repetitions, counterbalanced by referent position, of each Mandarin word with the English gloss of each of two Mandarin words differing minimally in lexical tone; 12 trials for each set) were identical to practice trials except that no feedback was provided. Responses were coded as correct, incorrect, or neither (e.g., if the child's pointing response was unclear or (s)he answered "I don't know"; 37 responses).

3 Results

For lexical and non-lexical tone discrimination, signal detection theory (Green & Swets, 1966; Macmillan & Creelman, 2004) was used to decompose responses into two conceptually and statistically distinct parameters: *discrimination* or *sensitivity* (*d′*), which captures how well participants successfully discriminated stimuli differing in tone from stimuli with the same tone (*d′*); and *response criterion* (*c*) or *response bias*, which captures the criterial level at which participants termed stimuli different, regardless of whether they actually differed. Data for these discrimination tasks were analyzed using mixed effects probability unit (probit) models, which operate on trial-level data and account for subject and item variability within the same model, obviating the data aggregation by these factors necessary for traditional analysis of variance (ANOVA). Unlike ANOVAs, probit mixed effects models allow participant responses (1 = same; 0 = different) rather than *d′* values to be used as the dependent variable (DV). Instead, measures of sensitivity derived from probit mixed effects models are expressed as *d′* values, promoting continuity with the results of traditional ANOVA signal detection analyses.

In the probit mixed effects models used to analyze the data, the relationship between the two stimuli (same = 1; different = -1) as well as language group (tonal-atonal bilinguals = 1; atonal-atonal bilinguals = -1) were included as fixed effects using the mean-centered (Helmert) contrast coding specified here. The intercept represents overall response bias (*c*), and the main effect of same versus different represents overall discrimination performance (*d′*), with an alpha level below 0.05 indicating that overall response bias and/or discrimination performance exceeded chance. The main effect of language group represents its effect on response bias (*c*), and the interaction of language group with stimulus relationship represents the effect of language group on discrimination performance (*d′*), with an alpha level below 0.05 indicating that the effect of language group on response bias and discrimination performance exceeded chance. In the model for non-lexical tone discrimination, items consisted of the tone of the second stimulus (in Hz), whereas in the model for lexical tone discrimination, items consisted of the second word (including lexical tone). Random slopes were included with the maximal random effect structure permitted to achieve model convergence (Barr, Levy, Scheepers, & Tily, 2013). Because all participants, regardless of language background, were presented with pairs of items that were either identical or differed, the models fit to the non-lexical and lexical tone discrimination data included random slopes for stimulus relationship by participant. Because language background differed between participants and because items were not hierarchically nested within stimulus relationship in these models, the final model included random intercepts for both participant and item (see Appendix 1 for models).

Word learning data were analyzed using a mixed effects logit model. This method models the log odds (or logit) of a correct response on each trial, with response correctness as the dependent variable (1 = correct; 0 = incorrect). In this model, language group (tonal bilinguals =1; atonal bilinguals = -1) was included as a fixed effect using the mean-centered (Helmert) contrast coding specified here. In this model, items consisted of the six words learned. Again, random slopes were included with the maximal random effect structure permitted to achieve model convergence (Barr et al., 2013). To account for individual differences in learning by participant and learnability by word, the final model included random intercepts for both participant and item (see Appendix 1 for model). All probit and logit mixed effects models were fit with Laplace estimation using the glmer() function of the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) using the BOBYQA optimizer to promote convergence in the R statistical programming language (R Core Team, 2015).

3.1 Non-lexical tone discrimination task

The distribution of level-pitch non-lexical tone discrimination by language background is displayed in Figure 3, and parameter estimates for the model are displayed in Table 3. In this task, discrimination accuracy was similar for tonal–atonal and atonal–atonal bilingual children. However, atonal–atonal bilingual children showed a conservative response bias, such that they

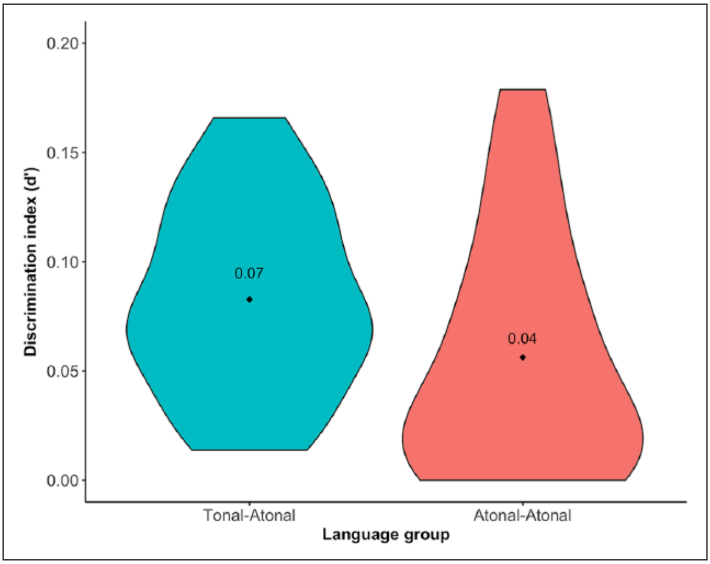

Figure 3. Non-lexical tone discrimination (*d′*) of tonal–atonal and atonal–atonal bilingual children (dots and values represent group means).

Table 3. Fixed effect estimates (top) and variance estimates (bottom) for multi-level probit model of non-lexical tone discrimination by language group (observations = 1120, log likelihood = -637.7).

Fixed effect	Coefficient	*SE*	Wald *z*	*p*
c (intercept)	0.22	0.25	0.85	.39
d′	0.49	0.25	1.97	.048
c × Language group	-0.25	0.10	-2.53	.01
d′ × Language group	0.11	0.08	1.28	.20
Random effect			s^2	
Participant			0.55	
Participant × *d′*			0.45	
Item			0.45	

were more likely to perceive sine wave tones differing in pitch as having the same pitch. These results indicate that tonal–atonal bilingual children are unable to discriminate between level-pitched sine wave tones any more accurately than atonal–atonal bilingual children. Furthermore, they suggest that experience with two atonal languages may desensitize children to small differences in F0, biasing them to perceive sine wave tones as similar even when they differ in pitch.

3.2 Lexical tone discrimination task

The distribution of lexical tone discrimination by language background is displayed in Figure 4, and parameter estimates for the model are displayed in Table 4. Consistent with predictions, tonal–atonal bilingual children's lexical tone discrimination accuracy (*d′*) was higher than that of

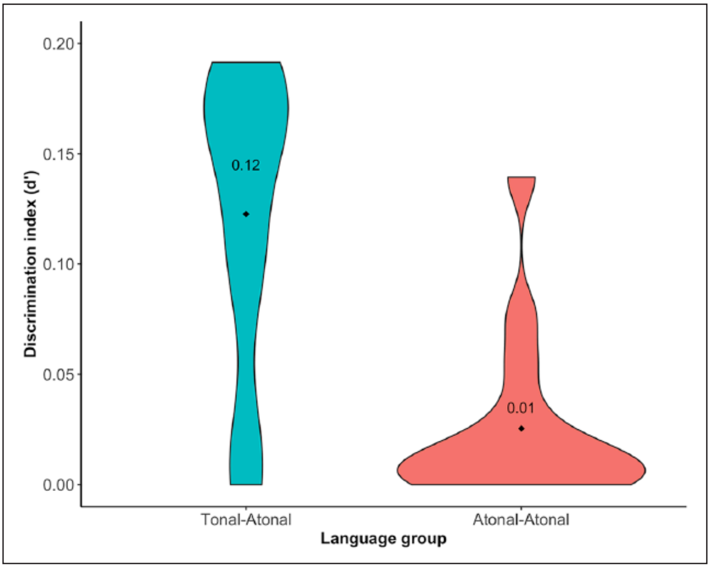

Figure 4. Lexical tone discrimination (*d′*) of tonal–atonal and atonal–atonal bilingual children (dots and values represent group means).

Table 4. Fixed effect estimates (top) and variance estimates (bottom) for multi-level probit model of lexical tone discrimination by language group (observations = 1440, log likelihood = -651.9).

Fixed effect	Coefficient	*SE*	Wald *z*	*p*
c (intercept)	0.64	0.23	2.72	.006
d′	0.25	0.27	0.92	.36
c × Language group	–1.87	0.36	–5.23	< .001
d′ × Language group	2.29	0.41	5.58	< .001
Random effect			s^2	
Participant			1.00	
Participant × *d′*			1.08	
Item			0.07	

atonal–atonal bilingual children. Again, atonal–atonal bilingual children showed a conservative response bias (*c*), such that they were more likely to perceive words differing minimally in lexical tone as having the same lexical tone. These results confirm that tonal–atonal bilingual children discriminate between words differing minimally in familiar lexical tone more accurately than atonal–atonal bilingual children. Furthermore, they suggest that experience with two atonal languages may desensitize children to differences in lexical tone, biasing them to perceive words differing minimally in lexical tone as similar even when they differ.

3.3 Word learning task

The distribution of word learning accuracy by language background is displayed in Figure 5, and parameter estimates for the model are displayed in Table 5. Consistent with predictions, the odds

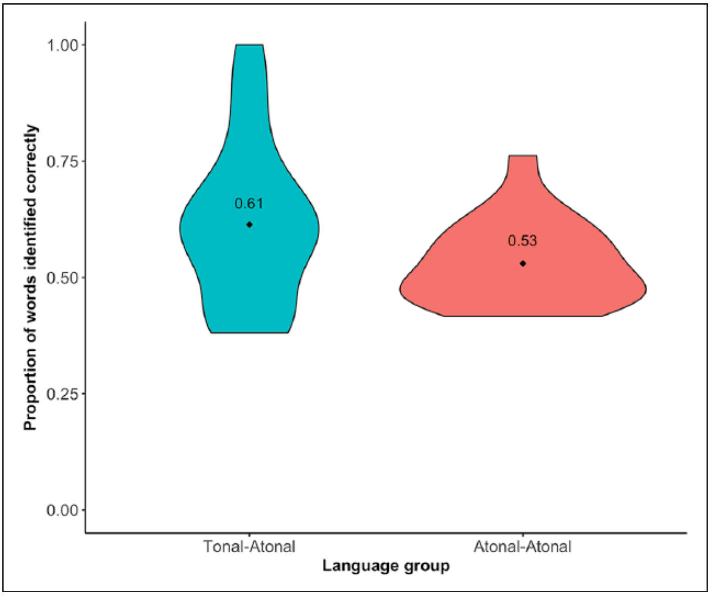

Figure 5. Proportion of words identified correctly in word learning task by tonal–atonal and atonal–atonal bilingual children (dots and values represent group means).

Table 5. Fixed effect estimates (top) and variance estimates (bottom) for multi-level logit model of word learning by language group (observations = 923, log likelihood = -625.7).

Fixed effect	Coefficient	*SE*	Wald *z*	*p*
Intercept	0.30	0.08	3.56	< .001
Language group	0.18	0.08	2.14	.03
Random effect			s^2	
Participant			0.32	
Item			< 0.01	

of correct identification of word referents was approximately 1.2 times (95% CI: 1.02, 1.42) greater for tonal–atonal bilinguals than for atonal–atonal bilinguals, indicating that tonal–atonal bilinguals had 20% greater odds than chance in selecting the correct referent of words differing minimally in lexical tone. These results provide evidence that experience with a tonal language in addition to an atonal language allows children to learn the meanings of novel words varying minimally in familiar lexical tone more effectively than experience with two atonal languages.

4 Discussion

The results of this work demonstrate that 5- to 7-year-old children bilingual in Mandarin, a tonal language, and English, an atonal language, perceive Mandarin lexical tone differently than children bilingual in English and another atonal language. With respect to lexical tone, the results demonstrate that, on average, Mandarin–English bilingual children are able to discriminate between Mandarin words differing minimally in lexical tone and map novel words differing minimally in Mandarin lexical tone onto their meanings more accurately than their atonal–English bilingual peers. Moreover, the results indicate that Mandarin–English and atonal–English bilingual children discriminate between level-pitched non-lexical tones with similar accuracy, suggesting that differences in Mandarin lexical tone perception do not extend to perception of these non-lexical tones. Together, these findings indicate that, relative to bilingualism in two tonal languages, bilingualism in a tonal language and an atonal language permits continued sensitivity to familiar lexical tone contrasts but does not enhance discrimination between level-pitched non-lexical tones beyond infancy.

The finding that 5- to 7-year-old tonal–atonal bilinguals discriminate between words containing familiar lexical tones more accurately than their atonal–atonal bilingual peers indicates that it is experience with a tonal language in particular, rather than bilingualism in general, that explains continued sensitivity to lexical tone beyond infancy. Thus, although in infancy, experience with multiple languages prolongs sensitivity to non-native phonemic and suprasegmental contrasts, including lexical tone (Graf Estes & Hay, 2015; Singh, 2018), this finding indicates that experience with multiple languages is insufficient to prolong this sensitivity beyond infancy. Moreover, this finding suggests that bilingualism does not attenuate sensitivity to phonemic contrasts present in only one language beyond infancy, despite work with younger children suggesting that it may (Singh et al., 2015; Wewalaarachchi et al., 2017). Indeed, Mandarin–English bilingual children's discrimination between Mandarin words differing minimally in lexical tone in the current study is as accurate as that of similarly-aged Thai monolingual children in previous work (Burnham et al., 1996). It is worth noting that, in the present study and that of Burnham and colleagues (1996), lexical tone discrimination was assessed by asking children to indicate whether words from their native tonal language differing minimally in lexical tone were the same or different. By contrast, work with younger children (Singh et al., 2015; Wewalaarachchi et al., 2017) has assessed lexical tone discrimination by examining sensitivity to lexical tone substitutions in words from children's native tonal language. Thus, lower sensitivity to lexical tone observed in this previous work may be explained by methodological differences in addition to—or in lieu of—age differences.

Similar to lexical tone discrimination, the finding that, relative to atonal–atonal bilingualism, tonal–atonal bilingualism enhances 5- to 7-year-old children's learning of novel words differing minimally in familiar lexical tone provides evidence that it is experience with a tonal language in particular, rather than bilingualism in general, that explains phonemic interpretation of lexical tone beyond infancy. This finding suggests that, beyond infancy, atonal–atonal bilinguals continue to interpret suprasegmental pitch—which forms the basis of lexical tone—non-phonemically, whereas tonal–atonal bilinguals continue to interpret it phonemically in the tonal language and non-phonemically in the atonal language, as has been shown at 0;11–0;13 (Singh & Foong, 2012; Singh et al., 2016). Although Mandarin monolingual children's learning of words differing minimally in Mandarin lexical tone was not examined in the current study, comparison of the results of current study to those of previous work with 2- to 3-year-old Mandarin monolinguals suggests that tonal monolinguals may be able to map words differing in familiar lexical tone onto their meanings even more effectively than tonal–atonal bilinguals (Yuet Sheung Lee et al., 2002). Such a pattern is consistent with the finding that similarly-aged Mandarin–English bilinguals are less sensitive than their Mandarin monolingual peers to lexical tone substitutions in Mandarin words (Wewalaarachchi et al., 2017). These findings highlight the difference between sensitivity to and phonemic interpretation of lexical tone, which may differ in tonal–atonal bilingual children relative to tonal monolingual children. However, the results of the current work show that tonal–atonal bilingual children are both more sensitive to acoustic differences between familiar lexical tones and more likely to interpret them phonemically than atonal–atonal bilinguals, demonstrating that experience with a tonal language in addition to an atonal language allows children to retain sensitivity to the acoustics and functions of familiar lexical tones, whereas experience with two atonal languages does not.

In contrast to lexical tone, tonal-atonal and atonal-atonal bilingual children discriminated between level-pitched sine wave tones with similar accuracy. This finding is consistent with the results of Burnham et al. (1996), which showed that 5- to 8-year-old monolingual Thai-speaking children, who discriminated between Thai lexical tones more accurately than their English-speaking peers, discriminated between sine wave tones and musical analogues of Thai lexical tones with accuracy comparable to that of English speakers. Thus, this finding suggests that, by the age of 5–7, experience with a tonal language in addition to an atonal language does not affect perception of level-pitched non-lexical tones. This finding is consistent with work demonstrating that adult monolingual tonal language speakers do not discriminate between level-pitched sine wave tones any more accurately than adult monolingual atonal language speakers (Bent et al., 2006; Burns & Sampat, 1980), but is inconsistent with other findings showing that these groups differ in level-pitched sine wave tone discrimination (Giuliano et al., 2011; Pfordresher & Brown, 2009; Stagray & Downs, 1993; Tanner & Rivette, 1964). In light of evidence suggesting that perception of non-lexical tones may vary according to their acoustic similarity to lexical tone, future research should examine how tonal–atonal and atonal–atonal bilingual children perceive contour-pitched analogues to familiar lexical tones. Such research would shed further light on the influences of tonal language experience and bilingualism on the development of non-lexical tone perception beyond infancy, clarifying whether such experience does not affect non-lexical tone perception at all or whether its effects are constrained to non-lexical tones that are acoustically similar to familiar lexical tones.

In addition to the discrimination patterns discussed above, the findings of the current work revealed that atonal-atonal bilingual children showed a conservative bias for both lexical and non-lexical tone discrimination, such that they tended to perceive differing tones as identical in pitch, whereas tonal-atonal bilingual children showed no such tendency. This finding suggests that a bias away from perceiving differences in pitch may arise from bilingualism or lack of experience with a tonal language, altering the perception of lexical and non-lexical tone. Future work should distinguish between these possibilities by comparing lexical and non-lexical tone discrimination in atonal monolinguals and atonal–atonal bilinguals. In general, these findings concerning perceptual biases suggest that bilingualism in a tonal language in addition to an atonal language may protect against the development of biases in lexical and non-lexical tone perception, suggesting that it may affect non-lexical tone perception in a subtle, limited manner.

The results of the present study represent an initial step towards filling three lacunae within the extant body of research on lexical and non-lexical tone perception: (a) the influence of tonal–atonal bilingualism on non-lexical tone perception; (b) whether atonal–atonal bilingualism impacts lexical tone perception similarly to tonal–atonal bilingualism; (c) the developmental trajectory of lexical and non-lexical tone perception beyond infancy. In particular, the present study allowed for comparison of the effects of experience with a tonal language in addition to an atonal language to the effects of bilingualism, regardless of language tonality, on lexical and non-lexical tone perception, which previous work had not considered. Through its examination of children aged 5;0–7;0 and use of experimental tasks similar to those used in previous work on lexical and non-lexical tone perception (e.g., Bradlow & Bent, 2008; Burnham et al., 1996; Yuet Sheung Lee et al., 2002), the current study permits a direct comparison of its results to the results of similar previous studies focusing on other age groups and/or language backgrounds. These comparisons reveal important information about lexical and non-lexical tone perception in tonal–atonal and atonal–atonal bilinguals that cannot be obtained from comparison of these groups with tonal and atonal monolinguals. For example, they indicate that tonal–atonal bilingualism permits better retention of familiar lexical tone discrimination beyond infancy than atonal–atonal bilingualism, and that neither bilingualism nor tonal language experience influences level-pitch sine wave tone discrimination beyond subtle, limited effects on bias. These findings suggest that by the age of 5;0, tonal–atonal bilinguals perceive familiar lexical tones phonemically and sine wave tones non-phonemically, whereas atonal–atonal bilinguals perceive both lexical and non-lexical sine-wave tones non-phonemically, highlighting the different mechanisms used by these populations to process lexical and non-lexical tone.

Despite the importance of these conclusions, they must be qualified by some important methodological caveats concerning the present study. First, Mandarin words differing minimally in lexical tone were used as stimuli in the lexical tone discrimination task. This may have provided tonal-atonal bilinguals with an advantage for words for which they knew the meaning, given that they could draw upon additional semantic cues to discriminate between them. Second, tonal-atonal bilinguals showed greater variation than atonal-atonal bilinguals in the proportion of words that they identified correctly in the word learning task, with some showing below-chance performance. Third, on a related note, the number of participants and trials used in this study may have produced results that were slightly underpowered, particularly given that a between-subjects design was necessary to examine the influence of language background on tone perception and word learning. Thus, this study should ideally be replicated with a larger (and possibly more varied) sample of tonal–atonal and atonal–atonal bilinguals to confirm the replicability and robustness of its findings.

Overall, the results of the present study indicate that bilingualism in a tonal language in addition to an atonal language allows for continued sensitivity to familiar lexical tone but does not enhance discrimination between level-pitched non-lexical tones beyond infancy. As such, the results provide empirical support for theories of phonological development in which experience with specific phonological features plays a primary role in shaping perceptual biases for speech sounds in infancy and beyond (e.g., Elman et al., 1997; Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996), particularly with regard to bilingualism (see Werker & Byers-Heinlein, 2008). Additionally, they suggest that tonal–atonal bilingualism may protect against development of biases to perceive differing lexical and level-pitched non-lexical tones as identical, thereby affecting non-lexical tone perception in a subtle, limited manner. Although it remains to be seen how bilingualism in languages with differing lexical tone systems influence lexical and non-lexical tone perception, the results of the present study demonstrate that tonal–atonal bilingualism—as well as atonal–atonal bilingualism—continues to influence perception of familiar lexical tones well beyond infancy.

Acknowledgements

The author thanks Kangya He for assistance with auditory stimulus preparation and data collection, Winston Goh, Leher Singh, Melvin Yap, and Brian MacWhinney for helpful discussion, and Tan Seok Hui for her support and help with all aspects of this project.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by an East Asia and Pacific Summer Institutes (EAPSI) Fellowship to the author (co-funded by the US National Science Foundation and the National Research Foundation of Singapore) and a grant from the Singapore Children's Foundation to Tan Seok Hui. The preparation of this manuscript was additionally supported by a National Defense Science and Engineering Graduate (NDSEG) Fellowship (32 CFR 168a) funded by the US Department of Defense, Air Force Office of Scientific Research), and by the Perlino Award of the Department of Psychology at UCSC, both to the author.

Notes

1. There is also a fifth tone, known as neutral or light tone, which lacks a distinct pitch contour; it is analogous to an unstressed syllable. This tone was not included in the current study due to its non-distinctive pitch contour, as well as its absence from some dialects of Mandarin and its rarity in others.
2. The results remained the same when these children were excluded from the analyses; thus, they are not discussed further, and the analyses reflect the results of all children.
3. The low dipping (3) tone, which is the least acoustically distinctive of the four principal Mandarin tones (Chao & Pian, 1948), was excluded to reduce the difficulty of the word learning task, which was quite challenging for children with experience only with atonal languages.

ORCID iD

Laura M. Morett https://orcid.org/0000-0002-1251-7213

References

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

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

Bent, T., Bradlow, A. R., & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 97–103.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011a). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, *23*, 425–434.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011b). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, *77*, 1–10.

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.

Blicher, D. L., Diehl, R. L., & Cohen, L. B. (1990). Effects of syllable duration on the perception of the Mandarin tone 2/tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics*, *18*, 27–49.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707–729.

Burnham, D., Francis, E., & Webster, D. (1996). The development of tone perception: Cross-linguistic aspects and the effect of linguistic context. In *Pan-Asiatic linguistics: Proceedings of the 4th International Symposium* (pp. 47–50). Salaya, Thailand: Institute of Language and Culture for Rural Development, Mahidol University. Retrieved from http://sealang.net/sala/archives/pdf4/burnham1996development.pdf

Burns, E. M., & Sampat, K. S. (1980). A note on possible culture-bound effects in frequency discrimination. *Journal of the Acoustical Society of America*, *68*, 1886–1888.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research*, *1128*, 148–156.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language*, *108*, 1–9.

Chao, Y. R. (1965). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.

Chao, Y. R., & Pian, R. C. (1948). *Mandarin primer (Vol. 26)*. Cambridge, MA: Harvard University Press.

Chen, A., & Kager, R. (2016). Discrimination of lexical tones in the first year of life. *Infant and Child Development*, *25*, 426–439.

Ciocca, V., & Lui, J. (2003). The development of the perception of Cantonese lexical tones. *Journal of Multilingual Communication Disorders*, *1*, 141–147.

Creel, S. C., Weng, M., Fu, G., Heyman, G. D., & Lee, K. (2018). Speaking a tone language enhances musical pitch perception in 3–5-year-olds. *Developmental Science*, *21*, e12503.

Delogu, F., Lampis, G., & Belardinelli, M. O. (2006). Music-to-language transfer effect: May melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive Processing*, *7*, 203–207.

Delogu, F., Lampis, G., & Belardinelli, M. O. (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology*, *22*, 46–61.

Deutsch, D., Dooley, K., Henthorn, T., & Head, B. (2009). Absolute pitch among students in an American music conservatory: Association with tone language fluency. *The Journal of the Acoustical Society of America*, *125*, 2398–2403.

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. S. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of the Acoustical Society of America*, *119*, 719.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1997). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Feng, A. (Ed.). (2007). *Bilingual education in China: Practices, policies and concepts*. Clevedon, UK: Multilingual Matters.

Francis, A. L., Ciocca, V., & Ng, B. K. C. (2003). On the (non) categorical perception of lexical tones. *Perception & Psychophysics*, *65*, 1029–1044.

Fromkin, V. A. (1978). *Tone: A linguistic survey*. New York: Academic Press.

Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, *11*, 149–175.

Gandour, J. (1984). Tone dissimiliarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, *12*, 235–261.

Gandour, J., Wong, D., & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *NeuroReport*, *9*, 2115–2119.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in Psychology*, *2*, 146.

Graf Estes, K., & Hay, J. F. (2015). Flexibility in bilingual infants' word learning. *Child Development*, *86*, 1371–1385.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hallé, P. A., Chang, Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, *32*, 395–421.

Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, *110*, 581–616.

Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child Development*, *86*, 10–22.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *97*, 3099–3111.

Hsieh, L., Gandour, J., Wong, D., & Hutchins, G. D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain and Language*, *76*, 227–252.

Hutka, S., Bidelman, G. M., & Moreno, S. (2015). Pitch expertise is not created equal: Cross-domain effects of musicianship and tone language experience on neural and behavioural discrimination of speech and music. *Neuropsychologia*, *71*, 52–63.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia*, *48*, 2630–2639.

Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage*, *13*, 646–653.

Kluender, K., Diehl, R., & Killeen, P. (1987). Japanese quail can learn phonetic categories. *Science*, *237*(4819), 1195–1197.

Kuhl, P. K. (1981). Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. *Journal of the Acoustical Society of America*, *70*, 340–349.

Lee, C.-Y., & Hung, T.-H. (2008). Identification of Mandarin tones by English-speaking musicians and non-musicians. *Journal of the Acoustical Society of America*, *124*, 3235–3248.

Lee, C.-Y., Yen, H., Yeh, P., Lin, W.-H., Cheng, Y.-Y., Tzeng, Y.-L., & Wu, H.-C. (2012). Mismatch responses to lexical tone, initial consonant, and vowel in Mandarin-speaking preschoolers. *Neuropsychologia*, *50*, 3228–3239.

Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, *25*, 527–542.

Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science*, *243*, 489–494.

Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., & Stewart, L. (2012). The mechanism of speech processing in congenital amusia: Evidence from Mandarin speakers. *PLoS One*, *7*, e30374.

Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, *133*, 385–394.

Liu, S., & Samuel, A. G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Language and Speech*, *47*, 109–138.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Mahwah, NJ: Lawrence Erlbaum Associates.

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*, 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*, 2032–2043.

Massaro, D. W., Cohen, M. M., & Tseng, C. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics*, *13*, 267–289.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, *10*, 241–265.

Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, *106*, 1367–1381.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.

Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain*, *133*, 2635–2642.

Oh, Y. M., Pellegrino, F., Coupé, C., & Marsico, E. (2013). Cross-language comparison of functional load for vowels, consonants, and tones. In F. Bimbot, C. Cerisara, C. Fougeron, G. Gravier, L. Lamel, F. Pellegrino, & P. Perrier (Eds.), *Proceedings of interspeech 2013* (pp. 3032–3036). Lyon, France: International Speech Communication Association.

Peng, G., Zheng, H.-Y., Gong, T., Yang, R.-X., Kong, J.-P., & Wang, W. S.-Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, *38*, 616–624.

Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics*, *71*, 1385–1398.

Quam, C., & Swingley, D. (2010). Phonological knowledge guides 2-year-olds' and adults' interpretation of salient pitch contours in word learning. *Journal of Memory and Language*, *62*, 135–150.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Singh, L. (2018). Bilingual infants demonstrate advantages in learning words in a third language. *Child Development*, *89*, 397–413.

Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, *124*, 128–142.

Singh, L., & Fu, C. S. (2016). A new view of language development: The acquisition of lexical tone. *Child Development*, *87*, 834–854.

Singh, L., Goh, H. H., & Wewalaarachchi, T. D. (2015). Spoken word recognition in early childhood: Comparative effects of vowel, consonant and lexical tone variation. *Cognition*, *142*, 1–11.

Singh, L., Hui, T. J., Chan, C., & Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science*, *17*, 94–109.

Singh, L., Poh, F. L. S., & Fu, C. S. L. (2016). Limits on monolingualism? A comparison of monolingual and bilingual infants' abilities to integrate lexical tone in novel word learning. *Frontiers in Psychology*, *7*, 667.

Singh, L., Tan, A., & Wewalaarachchi, T. D. (2017). Lexical tone variation and spoken word recognition in preschool children: Effects of perceptual salience. *Journal of Child Language*, *44*, 924–942.

Stagray, J. R., & Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics*, *21*, 143–163.

Surendran, D., & Levow, G. A. (2004). The functional load of tone in Mandarin is as high as that of vowels. In B. Bel & I. Marlien (Eds.), *Speech prosody 2004* (pp. 99–102). Nara, Japan: International Speech Communication Association.

Tanner, W. P., & Rivette, C. L. (1964). Experimental study of "tone deafness." *Journal of the Acoustical Society of America*, *36*, 1465–1467.

Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R. J., & Näätänen, R. (1999). Functional specialization of the human auditory cortex in processing phonetic and musical sounds: A magnetoencephalographic (MEG) study. *Neuroimage*, *9*, 330–336.

Tervaniemi, M., Medvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., Van Zuijen, T. L., & Näätänen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: A PET study. *Human Brain Mapping*, *10*, 74–79.

Tillmann, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., & Peretz, I. (2011). Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages. *Frontiers in Psychology*, *2*, 120.

Tompkins, C. A., & Flowers, C. R. (1985). Perception of emotional intonation by brain-damaged adults: The influence of task processing levels. *Journal of Speech, Language, and Hearing Research*, *28*, 527–538.

Tsao, F.-M. (2017). Perceptual improvement of lexical tones in infants: Effects of tone language experience. *Frontiers in Psychology*, *8*, 558.

Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, *78*, 332–348.

Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience*, *15*, 1019–1027.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, *106*, 3649–3658.

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Sciences*, *12*, 144–151.

Wewalaarachchi, T. D., Wong, L. H., & Singh, L. (2017). Vowels, consonants, and lexical tones: Sensitivity to phonological variation in monolingual Mandarin and bilingual English–Mandarin toddlers. *Journal of Experimental Child Psychology*, *159*, 16–33.

Wildgruber, D., Pihan, H., Ackermann, H., Erb, M., & Grodd, W. (2002). Dynamic brain activation during processing of emotional intonation: Influence of acoustic parameters, emotional valence, and sex. *Neuroimage*, *15*, 856–869.

Wong, P., & Leung, C. T.-T. (2018). Suprasegmental features are not acquired early: Perception and production of monosyllabic Cantonese lexical tones in 4-to 6-year-old preschool children. *Journal of Speech, Language, and Hearing Research*, *61*, 1070–1085.

Wong, P., Schwartz, R. G., & Jenkins, J. J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language, and Hearing Research*, *48*, 1065–1079.

Wong, P. C. M., Ciocca, V., Chan, A. H. D., Ha, L. Y. Y., Tan, L.-H., & Peretz, I. (2012). Effects of culture on musical pitch perception. *PLoS One*, *7*, e33424.

Wong, P. C. M., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping*, *28*, 995–1006.

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, *170*, 223–231.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics*, *27*, 55–105.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*, *120*, 1063–1074.

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, *68*, 123–139.
Yuet Sheung Lee, K., Chiu, S. N., & van Hasselt, C. A. (2002). Tone perception ability of Cantonese-speaking children. *Language and Speech*, *45*, 387–406.

Appendix 1

1.1 Probit mixed effects model fit to non-lexical tone discrimination data.

```
glmer(Response ~ 1 + Same_diff * Language_group + (1+Same_diff|Subject) + (1|Stim2), data=nonlexical_tone, family=binomial(link='probit'), control = glmerControl(optimizer="bobyqa"))
```

1.2 Probit mixed effects model fit to lexical tone discrimination data.

```
glmer(Response ~ 1 + Same_diff * Language_group + (1+Same_diff|Subject) + (1|Word2), data=lexical_tone1, family=binomial(link='probit'), control = glmerControl(optimizer="bobyqa"))
```

1.3 Logit mixed effects model fit to word learning data.

```
glmer(Correct ~ Language_group + (1|Subject) + (1|WordTone), data=word_learning, family=binomial, control = glmerControl(optimizer="bobyqa"))
```