Received: 27 January 2016 | Accepted: 13 August 2016
DOI: 10.1111/desc.12503

SHORT REPORT

WILEY Developmental Science

Speaking a tone language enhances musical pitch perception in 3–5-year-olds

Sarah C. Creel[1] | Mengxing Weng[2] | Genyue Fu[3] | Gail D. Heyman[4] | Kang Lee[2,4,5]

[1]Department of Cognitive Science, UC San Diego, La Jolla, CA, USA
[2]Zhejiang Normal University, Hangzhou, Zhejiang, China
[3]Hangzhou Normal University, Hangzhou, Zhejiang, China
[4]Department of Psychology, UC San Diego, La Jolla, CA, USA
[5]Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

Correspondence
Sarah Creel, UC San Diego, Department of Cognitive Science, 9500 Gilman Drive Mail Code 0515, La Jolla, CA 92093, USA.
Email: screel@ucsd.edu

Abstract

Young children learn multiple cognitive skills concurrently (e.g., language and music). Evidence is limited as to whether and how learning in one domain affects that in another during early development. Here we assessed whether exposure to a tone language benefits musical pitch processing among 3–5-year-old children. More specifically, we compared the pitch perception of Chinese children who spoke a tone language (i.e., Mandarin) with English-speaking American children. We found that Mandarin-speaking children were more advanced at pitch processing than English-speaking children but both groups performed similarly on a control music task (timbre discrimination). The findings support the Pitch Generalization Hypothesis that tone languages drive attention to pitch in nonlinguistic contexts, and suggest that language learning benefits aspects of music perception in early development. A video abstract of this article can be viewed at: https://youtu.be/UY0kpGpPNA0

RESEARCH HIGHLIGHTS

- Musical pitch perception is better in tone-language than non-tone-language children.
- This is far earlier (age 4 years) than previous findings of tone-language advantages.
- A control musical task (timbre) rules out overall better test performance.
- Findings imply strong perceptual permeability across domains.

1 | INTRODUCTION

One of the most fundamental questions about human development is the extent to which skills (or deficits) in one domain can lead to benefits (or costs) in other domains (e.g., Behrmann & Plaut, 2012; Hubbard, Piazza, Pinel, & Dehaene, 2005; Li et al., 2013). The answer to this question has far reaching theoretical implications for neuroscience and behavior, and has important practical implications for designing early-intervention programs or 'brain training' regimes. One way in which researchers have explored this question is by looking at relations between music skills and other kinds of cognitive skills in children. The focus of that research has been on whether music experience has implications for language development (e.g., Chobert, François, Velay, & Besson, 2014; François, Chobert, Besson, & Schön, 2013; Kraus et al., 2014; Moreno et al., 2011; see Patel, 2011, for a theoretical account).

In the present research we address the question of cross-domain learning by investigating potential influences in the opposite direction: whether language experience influences children's ability to make distinctions between musical sounds. In doing so we follow up on the hypothesis of Deutsch and colleagues (Deutsch, Henthorn, & Dolson, 2004; Deutsch, Henthorn, Marvin, & Xu, 2006; Henthorn & Deutsch, 2007) that experience with a tone language leads to enhanced pitch perception in music. However, unlike Deutsch and colleagues, we investigate perception of relative pitch rather than absolute pitch, and, for the first time, we address the issue from a developmental perspective. More specifically, we examine whether young children who have had experience with word-level linguistic pitch processing have an advantage in musical pitch processing, by comparing those who had exposure to a *tone language* to those without such exposure.

Tone languages (e.g., Mandarin, Thai, Yoruba, Xhosa) use pitch patterns on words to convey differences in meaning. For instance, in Mandarin, the syllable 'ma' can mean mother, horse, hemp, or scold depending on its pitch pattern. By testing tone-language-speaking and non-tone-language-speaking children, we can pit the *pitch specificity hypothesis*, in which pitch processing is specific to the context in which it is learned, against the *pitch generalization hypothesis*, in which pitch-processing advantages extend beyond that context.

While little research has directly examined whether or when during development pitch is processed differently depending on the context (music, tone language, non-tone language), several adult studies are consistent with the pitch specificity hypothesis. In general terms, Peretz and colleagues have argued for brain modularity of music processing (Peretz & Coltheart, 2003; Peretz & Zatorre, 2005). Speaking more specifically to pitch, Deutsch and colleagues have found adult behavioral (Deutsch, Henthorn, & Lapidis, 2011) and neural (Tierney, Dick, Deutsch, & Sereno, 2013) evidence for a more pitch-focused listening mode when hearing a spoken phrase as music vs. as a non-tone language (English). Burnham and colleagues (Burnham, Kasisopa et al., 2015) found that non-tone-language-speaking adults showed poor pitch-pattern discrimination for linguistic tones, but good discrimination for matched music-like sounds, suggesting that non-tone-language speakers' pitch processing differs depending on the context (language or music). Parallel findings exist in bilingual speech perception: Gonzales and Lotto (2013) found that English-like or Spanish-like phonetic context caused bilingual listeners to interpret the same speech sound as English *b* or Spanish *p*. If listeners can process sound patterns context-specifically for two languages, it stands to reason that they might process acoustic cues context-specifically for two more-dissimilar domains (i.e., language and music).

Other evidence supports the *pitch generalization hypothesis*, that attention to pitch in one context will benefit pitch processing in other contexts. Independent groups have reported a tone-language benefit for relative pitch processing in adults (Bidelman, Gandour, & Krishnan, 2011; Bidelman, Hutka, & Moreno, 2013; Hove, Sutherland, & Krumhansl, 2010; Hutka, Bidelman, & Moreno, 2015; Pfordresher & Brown, 2009; Wong et al. 2012; for absolute pitch, see Deutsch et al., 2006; Henthorn & Deutsch, 2007; although see Gregersen, Kowalsky, Kohn, & Marvin, 2000; and Schellenberg & Trehub, 2008, for differing perspectives and findings). In the opposite direction, among individuals without tone-language experience, musicians show better linguistic tone sensitivity compared to non-musicians (Burnham, Brooker, & Reid, 2015; Wong & Perrachione, 2007; Wong, Skoe, Russo, Dees, & Kraus, 2007; see also Patel, 2011). These data suggest that intensive exposure to consistent pitch patterns, via lifelong tone-language immersion or musical training, can improve pitch perception across domains.

However, it is not clear whether these adult findings point to actual overlap between systems, or simply adult-level cognitive and metalinguistic skills that permit mapping of one system onto the other. Few studies evaluate pitch specificity or pitch generalization hypotheses from a developmental perspective. Existing developmental evidence is equivocal, with one infant study supporting the pitch specificity hypothesis (Mattock & Burnham, 2006), but another study with school-aged children providing some support for the pitch generalization hypothesis (Peretz et al., 2013).

In the present study we directly tested these hypotheses in young children with a same/different discrimination task, which assesses the abilities to differentiate simple tone sequences by timbre vs. by pitch contour. In earlier work conducted in the United States, Creel (2014, 2016) found that children are better at distinguishing certain musical timbres than they are at distinguishing commonly encountered pitch contours. Thus, if tone language has no cross-domain effect on musical pitch processing, then both tone and non-tone groups should show an advantage for discriminating timbres. However, if tone language confers a benefit on musical pitch processing, Mandarin-speaking children should show better performance on pitch perception relative to English-speaking children. Different-timbre trials effectively serve as a task control.

2 | EXPERIMENT 1

2.1 | Method

2.1.1 | Participants

A target sample size of 48 was based on unpublished data from the Creel lab. Since more Mandarin-speaking children reached criterion (all of them; see Table 1), additional English-speaking children were tested to equate the number of criterion-reaching participants. In each group a few extra children were run to be inclusive of interested children. In the end, 51 Mandarin-speaking children (between 3.76 and 6.01 years, *M* = 4.86; 27 female) took part. None had had any private musical instruction. Children from the US (*N* = 53), ages between 3.68 and 5.66 years (*M* = 4.68; 31 female) also took part. An additional 23 American English-speaking children took part in the task but their data were replaced because they did not meet the training criterion (see below). Six further American English-speaking children met criterion but were excluded due to tone-language exposure (3), excessive noise (1), programming error (1), missing demographic data (1). For both Mandarin-speaking and English-speaking children, musical activities took place in their preschools.

TABLE 1 Participant characteristics

Exp.	Language group	Age, years (*SD*)	Age range	*N*	Met criterion	Accuracy
1	Non-tone (US English)	4.64 (0.47)	3.68–5.66	53	70%	0.747 (0.123)
	Tone (Mandarin Chinese)	4.85 (0.55)	3.76–6.01	51	100%	0.868 (0.101)
2	Under 4 years	3.85 (0.08)	3.71–3.997	42	81%	0.759 (0.121)
	Over 4 years	4.12 (0.07)	4.00–4.25	40	93%	0.823 (0.148)

Note. Age characteristics, *N*, and overall accuracy were calculated after eliminating participants who did not meet criterion or whose data were excluded.

2.1.2 | Stimuli

Two-tone sequences (duration: 400 milliseconds [ms] per tone plus 50 ms reverberation time, 850 ms total) were created using MIDI instruments in Finale software (2009, MakeMusic, Inc.). Files were exported as .aiffs to Praat software (Boersma & Weenink, 2014) to excise silences and normalize mean amplitude to 70 dB, then exported as .wav files for experimental presentation. In the examples below, numbers following note names indicate pitch height. Middle C is C4, and the B just below it is B3. The C an octave above middle C is C5.

Training stimuli were the two-note sequences B3-A4 and F4-E4, both in saxophone timbre. These pitch pairs differ in absolute pitch content (B and A, vs. F and E), contour (rising vs. falling), and interval size (10 semitones vs. 1 semitone), making them easy to distinguish along several pitch-related dimensions. Training stimuli were deliberately very distinct in order to provide children with clear, easily discernible examples of pitch differences. They differed in *pitch* rather than any other attribute, so that children were trained to pay attention to pitch attributes, the focus of the study.

There were three types of testing stimuli: timbre difference (for example, C4-G4 played by a trumpet vs. C4-G4 played by a vibraphone); contour difference (C4-G4 on trumpet vs. G4-C4 on trumpet); and timbre-order (C4[trumpet]-C4[vibraphone] vs. C4[vibraphone]-C4[trumpet]). The last trial type, timbre order, had been included in an earlier study to assess different hypotheses of negligible interest here; these trials are omitted from analyses. Within each test trial type, there were two pair types – timbre difference: trumpet vs. vibraphone, bassoon vs. saxophone; pitch contour: C4-G4 vs. G4-C4, C4-E4 vs. E4-C4; timbre order: trumpet-vibraphone vs. vibraphone-trumpet; bassoon-saxophone vs. saxophone-bassoon. The two pair types within each condition (such as C4-G4 vs. G4-C4 and C4-E4 vs. E4-C4) were included so as to present a range of differences in timbres and pitch contours.

Regarding pitch contour test stimuli, note that while children could succeed on training stimuli by using absolute pitch alone (without having to use pitch contour), pitch-contour *test* trials contained identical absolute pitches and were thus distinguishable by their contours (rising vs. falling) but not by absolute pitches alone. They *are* distinguishable by absolute pitch *order* (one starts with C, one with E). While one could tap relative pitch more directly by presenting stimuli that *always* change in absolute pitch, either retaining or changing contour (same trial: C4-G4 vs. E4-B4; different trial: C4-G4 vs. B4-E4), previous research (Bartlett & Dowling, 1980) suggests that this abstract relational task is too difficult for young children.

2.1.3 | Procedure

In the US, children were run in a quiet area in a preschool or day care facility on a Mac Mini running Matlab 2008a and Psychtoolbox3 (Brainard, 1997; Pelli, 1997). In China, children were run in a quiet room on a MacBook laptop running Matlab 2008a. In both locations, children wore child-sized KidzGear headphones.

There were three phases: pretraining, training, and test. During pretraining, children saw two visual examples (one different, one same) with feedback. Next was an auditory *same* example (B3-A4 followed by B3-A4) with a 1000-millisecond ISI, and children were asked to respond by saying 'same' or 'different'. If children did not respond 'same', the example was repeated up to three more times. Next, they heard a 'different' auditory example, with very-different pitches (B3-A4 followed by F4-E4). If children did not respond 'different', the 'different' auditory example was repeated up to three more times. After these examples, children continued to training.

During a block of randomly ordered training trials, children heard four different trials and four same trials (eight total), using training stimuli described above (different trials: B3-A4 and F4-E4, or F4-E4 and B3-A4; same trials: B3-A4 and B3-A4, or F4-E4 and F4-E4). Children responded to these trials and received feedback ('Good job!' or 'No, those two were the same/different.') If they did not meet the training criterion of answering correctly on at least seven of eight trials in a block, the block was repeated, up to four more times (total of five). Children whose data did not meet criterion (at least 7/8 correct) after five blocks were excluded from analyses. Note that because training trials required children to detect a pitch difference, and previous studies on children in the US (Creel, 2014, 2016) have shown that pitch differences are somewhat difficult for children to detect, training trials might be more difficult for English-speaking children.

During test, trials were presented much as during training, except without feedback and with a larger variety of stimuli. There were eight 'different' trials and eight 'same' trials for each condition (contour, timbre, timbre order). To assess continued task adherence, eight trials with the training stimuli were also included (four same, four different). Thus the test contained 56 total trials.

2.2 | Results

Accuracy scores[1] were entered into an ANOVA with Language Background (Mandarin Chinese, US English) as a between-subjects factor and Trial Type (contour, timbre) as a within-subjects factor. Critically, if there are differences in performance as a function of language background, there should be an interaction of Language Background × Trial Type.

Overall, Mandarin-speaking children were more accurate than English-speaking children (Figure 1), reflected in a main effect of Language Background ($F(1, 102) = 21.68$, $p < .0001$, $\eta^2_p = .18$). A main effect of Trial Type ($F(1, 102) = 16.22$, $p = .0001$, $\eta^2_p = .14$) reflected overall higher accuracy on timbre trials than pitch contour trials. Finally, there was an interaction of Language Background × Trial Type ($F(1, 102) = 13.85$, $p = .0003$, $\eta^2_p = .12$). Specifically, tone-language speakers outperformed non-tone-language speakers on

[1]Analyses in both experiments showed identical significance patterns when converted to *d*-prime scores, using the independent-observations formula (MacMillan & Creelman, 2005).

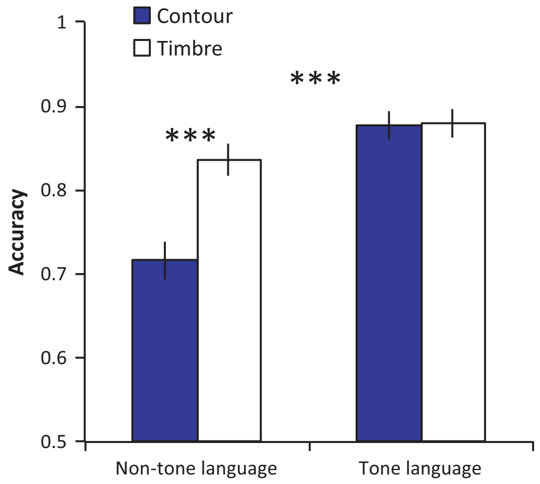

FIGURE 1 Experiment 1, pitch contour trials vs. timbre trials, with standard errors. Upper set of asterisks refers to Language Background × Trial Type interaction. ***p < .001

pitch trials (t(102) = 5.73, p < .0001, 95% CI: [.105, .216], d = 1.13). For timbre trials, the numerical advantage for tone-language speakers missed significance (t(102) = 1.69, p = .09, 95% CI: [−.007, .095], d = 0.33).

2.3 | Discussion

We found that young (3–5-year-old) tone-language speakers excel at a musical pitch perception task when compared to non-tone-language speakers. This is consistent with the pitch generalization hypothesis that language experience should affect pitch processing in music, and is inconsistent with the pitch specificity hypothesis that linguistic and musical pitch processing are largely separate. However, because Mandarin-speaking participants performed better overall, it is possible that our findings are due to near-ceiling performance in all trial types for Mandarin speakers (though only five Mandarin speakers, 10%, scored perfectly). Further, fewer of the English-speaking children met the training criterion (χ^2 = 16.86. p < .0001). This may reflect that Mandarin-speaking children were more successful than English-speaking children on training trials because those trials required pitch discrimination, consistent with the hypothesis that tone languages confer pitch-processing advantages. Alternatively, it may mean that the task was *overall* harder for English-speaking children. To address these issues, we conducted Experiment 2.

3 | EXPERIMENT 2

Here, we aimed to replicate the findings of Experiment 1 while matching tone- and non-tone-language-speaking children for overall performance level. We did this by testing younger Mandarin-speaking children. If musical pitch processing *is* driven by language, as suggested by the pitch generalization hypothesis, then younger Mandarin speakers should continue to show better pitch processing than English-speaking children in Experiment 1.

3.1 | Method

3.1.1 | Participants

We tested N = 82 children in the age range 3.7–4.3 years (M = 3.98, SD = .16, 42 female), younger than the Mandarin-speaking sample in Experiment 1 (M = 4.85 years). The age of this sample was informed by additional data collected from young Mandarin-speaking 3-year-olds (2.7–3.4 years), very few of whom managed to meet the training criterion (15 of 60, or 25%). An additional 13 children were tested but did not meet the training criterion. Again, the target sample size was 48. However, a miscommunication resulted in collection of a nearly complete duplicate sample (total N = 95; 82 met criterion). We elected to split the sample by age into two approximately equally sized groups.

Of children whose parents returned surveys (39), 9 had some sort of music instruction on pitched instruments (5 on specific instruments; 4 were in classes based on the Orff system of music pedagogy, a play-based system which exposes children to music through playing pitched and unpitched percussion, dancing, and drama; http://aosa.org/about/what-is-orff-schulwerk/). Analyses suggested that children with music lessons did not show heightened performance relative to children without lessons on pitched musical instruments, nor did music training interact with trial type (pitch contour vs. timbre). Therefore we collapsed analyses across music experience.

3.1.2 | Stimuli and procedure

These matched Experiment 1.

3.2 | Results

We examined Mandarin-speaking children's performance on the training trials and found their success on training trials differed depending on children's age: whereas only 7% of children age 4 or older failed to meet criterion (chi-square test vs. US children: χ^2 = 7.41, p = .006), children under age 4 years were less likely to meet criterion (nearly 20% did not). The younger group's performance was not significantly different from that of the 3–5-year-old (M = 4.64) US children (χ^2 = 1.43, p = .23). We therefore divided the children in Experiment 2 into two age groups (younger than 4 years and older than 4 years) and compared the Experiment 1 English-speaking sample to each of these groups individually.

First, we compared the younger Mandarin-speaking children (N = 42, 19 female, M = 3.85 years) to English-speaking children in Experiment 1. Here, English-speaking and Mandarin-speaking samples were equivalent in overall accuracy (no main effect of Language Background; F(1, 93) = 0.02, p = .89, η^2_p = .00). As before, timbre trials showed overall higher accuracy (Trial Type, F(1, 93) = 14.01, p = .0003, η^2_p = .13). As previously, there was a Language Background × Trial Type interaction (F(1, 93) = 10.16, p = .002, η^2_p = .10). The interaction resulted from tone-language speakers showing a marginal advantage on contour trials (t(93) = 1.82, p = .07, 95% CI: [−.01–.13], d = 0.37) while non-tone-language speakers had a marginal advantage on timbre trials ($t(93) = 1.82$, $p = .07$, 95% CI: [−.005–.11], $d = 0.37$). Thus, with equivalent overall accuracy, the two groups showed different relative advantages.

Second, we compared the children from the current sample aged 4 years or older ($N = 40$, 23 female, age $M = 4.12$ years; Figure 2) to the English-speaking children in Experiment 1. We found that the Mandarin-speaking older children were significantly more accurate than English-speaking children overall (main effect of Language Background, $F(1, 91) = 5.29$, $p = .02$, $\eta^2_p = .05$). There was also an effect of Trial Type ($F(1, 91) = 14.65$, $p = .0002$, $\eta^2_p = .14$), and a Language Background × Trial Type interaction ($F(1, 91) = 17.36$, $p < .0001$, $\eta^2_p = .16$). Significantly, this group of Mandarin-speaking children was the one best matched to English speakers in timbre accuracy ($t(91) = 0.16$, $p = .87$, 95% CI: [−.056, .065], $d = .03$), yet still showed a large advantage for pitch trials ($t(91) = 3.94$, $p = .0002$, CI: [.065, .197], $d = .83$).

3.3 | Discussion

Younger Mandarin-speaking children (ages 3.7–4.3 years) showed lower overall accuracy and lower rates of meeting the training criterion than their older counterparts from Experiment 1, but nonetheless showed relative advantages in detecting pitch-contour changes. This pattern is very similar to the older Mandarin-speaking children from Experiment 1, and very unlike non-tone-language-speaking children. This replicates and reinforces our finding in Experiment 1 that tone-language-speaking children have stronger pitch perception skills than non-tone-language-speaking children.

4 | GENERAL DISCUSSION

Young children who spoke a tone language showed an advantage in pitch contour perception compared to children who spoke a non-tone

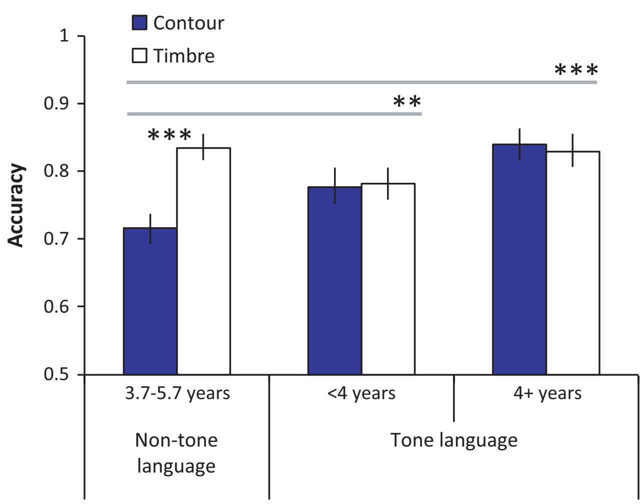

FIGURE 2 Experiment 2, later samples of tone-language speakers, with standard errors. Upper two sets of asterisks refer to Language Background × Trial Type interaction. ** $p < .005$; *** $p < .001$

language. The latter group, consistent with earlier studies (Creel, 2014, 2016), showed a disadvantage in discriminating pitch contours relative to timbres. Mandarin-speaking children, however, showed pitch contour discrimination that equaled their timbre discrimination abilities. Crucially, results held when we equated for task performance by testing younger Mandarin-speaking children.

Our findings are consistent with the *pitch generalization hypothesis*, providing evidence that tone languages draw attention to pitch in the language domain, which then confers benefits in musical pitch processing. Further, we find these pitch-processing advantages in children under 4 years, substantially earlier than any previous findings of tone-language benefits. Since none of the children had received literacy instruction to denote lexical tone (which begins after 6 years of age in China), our finding further suggests that tone-language benefits on musical pitch processing are not a result of literacy instruction. This is critically important in that existing research suggests that overt awareness of lexical tone in tone languages may be facilitated by tone cues in the language's writing system(s) (Burnham et al., 2011). Thus, unlike previous findings of facilitated pitch processing with adults, our pitch advantage findings cannot be explained as a side effect of orthographic instruction drawing attention to pitch.

Our data represent the earliest developmental time point where tone-language exposure influences non-speech pitch processing, implying that as few as four years of tone-language exposure may strengthen pitch processing in nonlinguistic sounds. Together with Mattock and Burnham (2006), who failed to find tone-language-based musical pitch-processing advantages in 9-month-olds, and adult studies that found tone-language-based advantages in musical pitch processing (Bidelman et al., 2011, 2013; Pfordresher & Brown, 2009; Wong et al. 2012), our findings suggest that tone-language effects on pitch processing may emerge after infancy, in early childhood.

Our findings are inconsistent with those of Peretz et al. (2013). They tested older children than we did (6–8 years) and came to a different conclusion – that Mandarin-speaking children did not possess a specific pitch-processing advantage, but an overall task performance advantage, over non-tone-language-speaking children. One reason may be that, while we, like Peretz et al. (2013), found better overall task performance in Mandarin-speaking children, we also tested younger Mandarin-speaking children who showed weaker task performance, and differences in pitch contour processing persisted. Second, Peretz and colleagues' music aptitude measure did not include a test for timbre perception, which served as a valuable control in our study.

Although we attribute the advanced pitch-processing skills of Mandarin-speaking children to tone-language experience, other explanations remain. We cannot rule out a possible role for genetic factors (Gregersen et al., 2000; Hove et al., 2010; see also Dediu & Ladd, 2007, on possible genetic loci implicated in linguistic pitch perception). Future work might consider testing children of East Asian ethnicity who are being raised outside Asia and learning a non-tone language; or testing children of non-Asian ethnicity outside Asia who are learning a tone language (e.g., an African tone language). Future research is also needed to determine the extent of the pitch-processing advantage among tone-language speakers. Given that speakers of tone languages are better at distinguishing lexical tones in their own language than tones in unfamiliar languages (Burnham, Kasisopa et al., 2015), it seems possible that pitch facilitation is strongest for pitch patterns that map onto native tones. Nor can we rule out additional cognitive or cultural differences between American and Chinese children, including differences in executive function (Sabbagh, Xu, Carlson, Moses, & Lee, 2006), educational achievement (Geary, 1996; Stevenson, Chen, & Lee, 1993), or both.

More broadly, our findings suggest that Deutsch (Deutsch et al., 2004, 2006) was correct that tone languages facilitate musical pitch processing. Our finding that language exposure changes *relative* pitch processing makes it plausible that similar exposure-based cross-modal influences might alter absolute pitch processing. Of course, our pitch-contour stimuli are also distinguishable by absolute pitch order. This leaves open the possibility that Mandarin-speaking children are better at absolute pitch rather than, or in addition to, relative pitch. We regard this as an interesting interpretation because there is as yet no empirical evidence for widespread absolute pitch perception in young tone-language-speaking children. Future work is needed, perhaps using different dependent measures, to ascertain what aspects of pitch processing are advantaged in Mandarin-speaking groups.

Effects of language on musical pitch perception support findings of cross-domain behavioral and neural interactions between language and music in adults. Previous studies on adult tone-language speakers suggest that tone languages boost musical pitch perception relative to non-tone-language speakers (e.g., Bidelman et al., 2011, 2013; Pfordresher & Brown, 2009). In the opposite direction, a number of studies also suggest that musical experience enhances encoding of linguistic pitch (Burnham, Brooker et al., 2015; Wong et al., 2007) and learning of new languages with tone content (Wong & Perrachione, 2007) in adults, and enhances language processing in children (e.g., Chobert et al., 2014; François et al., 2013; Kraus et al., 2014; Moreno et al., 2011; though see Mehr, Schachner, Katz, & Spelke, 2013). Our findings, together with this previous body of work, suggest that there is substantial permeability between linguistic and musical domains. While Patel's (2011) OPERA hypothesis suggests that these effects may be unidirectional (music influences linguistic pitch processing), we find strong evidence for the opposite direction of influence (linguistic pitch content influences musical pitch processing).

Given the robustness of our results, we suggest that this cross-domain permeability may be especially strong in early childhood. Perceptual-cognitive permeability lends itself to possibilities for early remediation of language processing difficulties, and implies a complex behavioral and neural interplay between non-modular cognitive capacities.

In summary, we show for the first time that tone-language experience is associated with advanced musical pitch processing in young children. This represents by far the earliest time point at which such an effect has been documented, and suggests that tone-language exposure facilitates pitch processing even in the absence of metalinguistic tone awareness. Results support the pitch generalization hypothesis and imply substantial cross-domain permeability between music and language.

ACKNOWLEDGEMENTS

SCC was supported by NSF CAREER Award BCS-1057080. KL and GF were supported by National Natural Science Foundation of China Grants 31371041 and 31470993, and NSERC. In loving memory of Laura Creel and Liesl Creel.

REFERENCES

Bartlett, J.C., & Dowling, W.J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 501–515.

Behrmann, M., & Plaut, D.C. (2014). Bilateral hemispheric processing of words and faces: Evidence from word impairments in prosopagnosia and face impairments in pure alexia. *Cerebral Cortex*, *24*, 1102–1118.

Bidelman, G.M., Gandour, J.T., & Krishnan, A. (2011). 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.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer [Computer program]. Version 5.4.01, retrieved 9 November 2014 from http://www.praat.org/

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Burnham, D., Brooker, R., & Reid, A. (2015a). The effects of absolute pitch ability and musical training on lexical tone perception. *Psychology of Music*, *43*, 881–897.

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., … Webster, D. (2015b). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics*, *36*, 1459–1491.

Burnham, D., Kim, J., Davis, C., Ciocca, V., Schoknecht, C., Kasisopa, B., & Luksaneeyanawin, S. (2011). Are tones phones? *Journal of Experimental Child Psychology*, *108*, 693–712.

Chobert, J., François, C., Velay, J.-L., & Besson, M. (2014). Twelve months of active musical training in 8- to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*, 956–967.

Creel, S.C. (2014). Tipping the scales: Auditory cue weighting changes over development. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1146–1160.

Creel, S.C. (2016). Ups and downs in auditory development: Preschoolers' sensitivity to pitch contour and timbre. *Cognitive Science*, *40*, 373–403.

Dediu, D., & Ladd, D.R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences, USA*, *104*, 10944–10949.

Deutsch, D., Henthorn, T., & Dolson, M. (2004). Absolute pitch, speech, and tone language: Some experiments and a proposed framework. *Music Perception*, *21*, 339–356.

Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *Journal of the Acoustical Society of America*, *129*, 2245–2252.

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (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*, 719–722.

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*, 2038–2043.

Geary, D.C. (1996). International differences in mathematical achievement: Their nature, causes, and consequences. *Current Directions in Psychological Science*, *5*, 133–137.

Gonzales, K., & Lotto, A.J. (2013). A bafri, un pafri: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science*, *24*, 2135–2142.

Gregersen, P.K., Kowalsky, E., Kohn, N., & Marvin, E.W. (2000). Letter to the editor: Early childhood music education and predisposition to absolute pitch: Teasing apart genes and environment. *American Journal of Human Genetics*, *98*, 280–282.

Henthorn, T., & Deutsch, D. (2007). Ethnicity versus early environment: Comment on 'Early childhood music education and predisposition to absolute pitch: Teasing apart genes and environment'. *American Journal of Medical Genetics*, *282*, 102–103.

Hove, M.J., Sutherland, M.E., & Krumhansl, C.L. (2010). Ethnicity effects in relative pitch. *Psychonomic Bulletin & Review*, *17*, 310–316.

Hubbard, E.M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, *6*, 435–448.

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.

Kraus, N., Slater, J., Thompson, E.C., Hornickel, J., Strait, D.L., Nicol, T., & White-Schwoch, T. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *Journal of Neuroscience*, *34*, 11913–11918.

Li, S., Lee, K., Zhao, J., Yang, Z., He, S., & Weng, X. (2013). Neural competition as a developmental process: Early hemispheric specialization for word processing delays specialization for face processing. *Neuropsychologia*, *51*, 950–959.

MacMillan, N.A., & Creelman, C.D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Taylor & Francis.

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

Mehr, S.A., Schachner, A., Katz, R.C., & Spelke, E.S. (2013). Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment. *PLoS ONE*, *8*, e82007.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E.G., Cepeda, N.J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, *22*, 1425–1433.

Patel, A.D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis *Frontiers in Psychology*, *2*, 142.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688–691.

Peretz, I., Gosselin, N., Nan, Y., Caron-Caplette, E., Trehub, S.E., & Béland, R. (2013). A novel tool for evaluating children's musical abilities across age and culture. *Frontiers in Systems Neuroscience*, *7*, 30.

Peretz, I., & Zatorre, R.J. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*, 89–114.

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

Sabbagh, M.A., Xu, F., Carlson, S.M., Moses, L.J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and US preschoolers. *Psychological Science*, *17*, 74–81.

Schellenberg, E.G., & Trehub, S.E. (2008). Is there an Asian advantage for pitch memory? *Music Perception*, *25*, 241–252.

Stevenson, H.W., Chen, C., & Lee, S.Y. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science*, *259*, 53–58.

Tierney, A.T., Dick, F., Deutsch, D., & Sereno, M. (2013). Speech versus song: Multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. *Cerebral Cortex*, *23*, 249–254.

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. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*, 565–585.

Wong, P.C.M., Skoe, E., Russo, N.M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.