

Simultaneous bilinguals who do not speak a tone language show enhancement in pitch sensitivity but not in executive function

Liquan Liu,^{1,2,3} Ao Chen,^{3,4} and René Kager³

¹ Western Sydney University | ² University of Oslo | ³ Utrecht University |

⁴ Beijing Language and Culture University

Previous studies have reported perceptual advantages, such as when discriminating non-native linguistic or musical pitch differences, among first-year infants growing up in bilingual over monolingual environments. It is unclear whether such effects should be attributed to bilinguals' enhanced perceptual sensitivity and/or cognitive abilities, and whether such effects would extend to adulthood. Twenty-four Dutch, 24 Dutch simultaneous bilingual (DSB), and 24 Chinese Mandarin speakers were examined by three sets of tasks assessing their linguistic pitch and music perception, executive function, as well as interactions across these modalities. Results showed degrees of advantages for DSB and Chinese participants' over their Dutch peers in lexical tone discrimination and pitch-related music tasks. In tasks related to executive function, no difference was observed between DSB and Dutch participants, while Chinese participants' performances were modulated by cognitive interference of language processing. Findings suggest that listeners' enhanced sensitivity to linguistic and musical pitch may stem from acoustic (DSB) and experience (Chinese) rather than cognitive factors. Moreover, Dutch participants showed robust correlations between their linguistic and musical pitch perception, followed by limited correlations in DSB, and virtually no correlation among Chinese participants, illustrating how distinct language experiences can lead to specific pitch perception patterns between language and music.

Keywords: lexical tone, musical pitch, bilingualism, enhanced sensitivity, executive function

Language and music are two universal human capacities that coexist across cultures and societies (Nettl, 2000; Patel, 2008). The exploration of the inter-relationship between the two domains, such as the perception of linguistic and

musical pitch, can increase our understanding of domain-general and -specific mechanisms that drive auditory perception. Meanwhile, listeners' linguistic experience plays a key role in their auditory perception and arguably cognitive functions. With respect to auditory perception, infants growing up in a bilingual environment show enhanced sensitivity to linguistic and musical tones (Liu & Kager, 2017a, 2017b). It remains unclear whether such sensitivity can be observed in bilingual adults as well as bilingual children. In addition, bilingualism has also been argued to lead to enhanced inhibitory control and executive functions, among other cognitive abilities (Bialystok, Craik, Klein & Viswanathan, 2004). This study was conducted to examine the inter-relationship among listeners' linguistic experience, pitch perception and their executive function.

Human sensitivity to prosodic features in language and music (e.g., pitch, rhythm) emerges in neonates (Nazzi, Bertoncini & Mehler, 1998; Winkler, Háden, Ladinig, Sziller & Honing, 2009). In the first year after birth, infants attune to their native language (Werker & Tees, 2002) and culture-specific music conventions (Hannon & Trainor, 2007; Soley & Hannon, 2010). In the speech domain, this process affects our perception of not only segmental (Kuhl et al., 2006) but also suprasegmental features (Mattock & Burnham, 2006) such as lexical tones, word-level pitch variations that lead to changes in meaning (Yip, 2002). Lacking lexical tonal experience, post-attunement infants (Liu & Kager, 2014) and adults (Song, Skoe, Wong & Kraus, 2008) exposed to non-tone languages do not discriminate tones in the same way as their tone-language peers.

Much debate has been given to pitch lateralization in the processing of linguistic pitch. The functional hypothesis argues for an asymmetry in pitch processing between cerebral hemispheres based on functions of pitch. On the one hand, a left-hemispheric dominance would surface for pitch patterns carrying heavier linguistic load like lexical tone whereas a right-hemispheric dominance would appear for those with lighter linguistic load such as affective intonation. Such asymmetry appears to be exaggerated in tone language speakers but not in non-tone language listeners (Bidelman & Chung, 2015; Wang, Jongman & Sereno, 2001). On the other hand, the acoustic hypothesis argues for a right-hemispheric dominance for general pitch processing (Wong, 2002). Irrespective of the competing hypotheses for linguistic pitch, a right-hemispheric superiority is reported in musical pitch perception (Francis, Ciocca, Ma & Fenn, 2008; Klein, Zatorre, Milner & Zhao, 2001; Tervaniemi et al., 1999; Zatorre, Belin & Penhune, 2002). Taken together, studies suggest asymmetrical perception between linguistic and musical pitch (Gandour et al., 2004).

In addition, neurological and psychophysiological studies demonstrate parallel resources at neural levels across the two domains (Koelsch et al., 2002; Patel, 2008). The memory and learning mechanisms for language and music, as well

as sound, prosodic, and grammatical structures of the two fields share similar developmental underpinnings (McMullen & Saffran, 2004). Compared to their non-tone language peers, tone language speakers show better imitation and perceptual discrimination ability of musical pitch (Pfordresher & Brown, 2009). The enhanced musical pitch perception and production among tone-language listeners may stem from their improved sensitivity in detecting subtle pitch contour and interval changes. A similar correlation has been found in the opposite direction, where musicians outperform their non-musician peers in not only musical (Burnham & Brooker, 2002) but also linguistic pitch perception (Bidelman, Hutka & Moreno, 2013; Delogu, Lampis & Olivetti Belardinelli, 2006, 2010; Lee & Hung, 2008). Musical training appears to sharpen listeners' subcortical encoding of general pitch patterns (Patel & Iversen, 2007; Wong, Skoe, Russo, Dees & Kraus, 2007). Auditory brainstem response studies reveal that both tone language and music experience can increase listeners' linguistic and musical pitch-tracking accuracy (Bidelman, Gandour & Krishnan, 2011a). Moreover, non-tone language listeners exhibit a correlation of their perceptibility between linguistic and music pitch (Chen, Liu & Kager, 2016): Speakers of Chinese and Dutch, a tone and non-tone language respectively, were tested on their linguistic pitch discrimination using the Montreal Battery of Evaluation of Amusia (MBEA, Peretz, Champod & Hyde, 2003) and Musical Ear Test (Wallentin, Nielsen, Friis-Olivarius, Vuust & Vuust, 2010). Chinese participants outperformed their Dutch peers on the two music test results. In addition, a correlation was found between non-tone language listeners' linguistic pitch and musical test performances, a pattern missing among tone language listeners. The findings are two-fold. On the one hand, shared neural resources may be recruited for the processing of language and music, and listeners' fine-tuning of critical auditory dimensions in one domain may impact their ability in the other. On the other hand, results favour a unified processing of cross-domain pitch perception for non-tone language listeners, but for tone language listeners, lexical tones are perceived as phonological categories, different from musical tones which do not play a phonemic role. Evidence of a split hypothesis also emerges in amusia studies where tone language speakers show successful production and discrimination of native lexical tones (Liu et al., 2012; Nan, Sun & Peretz, 2010).

Given that one's linguistic experience influences one's perception of language and music, it is interesting to examine the potential impact of bilingualism on speech and music processing. Bilinguals have been shown to exhibit enhanced sensitivity and perceptual flexibility to phonetic distinctions (Werker, 1986), which subsequently manifest in their advantage in learning non-native speech contrasts phonetically similar to their native language(s) (Antoniou, Liang, Ettliger & Wong, 2015). In addition, their subcortical encoding of fundamental

frequency (Fo), a feature underlying pitch perception and grouping of auditory objects, appears to be enhanced (Krizman, Marian, Shook, Skoe & Kraus, 2012). These advantages emerge from infancy, where bilingual infants demonstrate enhanced sensitivity in language and music perception (Liu & Kager, 2016). For example, while Dutch infants reached adult-like sensitivity to linguistic pitch by 18 months, simultaneous-bilingual Dutch infants achieved approximately the same level of sensitivity 6 months earlier (Liu & Kager, 2014, 2017a). Moreover, when hearing a violin contrast that differed solely in Fo, 9-month-old Dutch bilingual infants showed more robust perceptual sensitivity than their monolingual peers (Liu & Kager, 2017b). Studies have reported that listeners' auditory system encodes Fo more robustly and efficiently if they grow up in a bilingual environment, promoting experience-dependent neural plasticity (Krizman et al., 2012). These findings lead to the hypothesis that bilinguals have enhanced auditory sensitivity.

The nature of such enhanced sensitivity remains a matter of discussion. One explanation is that such an advantage may stem from the complex linguistic environment bilinguals are facing. The Native Language Magnet Theory Expanded (NLM-e) model (Kuhl et al., 2008) suggests that more condensed phonetic environments may lead to more specialized category formation during the neural commitment stage given sufficient input. It is not unlikely that bilinguals, having refined categories from both of their languages, would transfer this to general auditory perception, evidenced from abovementioned findings.

Another possibility is that bilinguals may have better cognitive control abilities. After all, auditory perception cannot be singled out from other cognitive functions such as attention and working memory. The effect of bilingualism on cognitive functions has been widely studied across the lifespan (Bialystok, Martin & Viswanathan, 2005; Carlson & Meltzoff, 2008). It has been argued that bilinguals need to constantly select the right language and suppress the irrelevant ones based on the given context or situation, which fine-tunes their cognitive control capacities that require selective attention and inhibition such as executive function. Executive function is an umbrella term that incorporates a collection of inter-related processes responsible for goal-directed behavior (Anderson, 2002; Gioia, Isquith & Guy, 2001). For example, it has been argued that bilinguals' constant code-switching of their two languages will strengthen their abilities to attend to relevant aspects of a situation and ignore irrelevant information (Bialystok, 1999). The bilingual environment appears to yield structural and functional changes in cortical and subcortical regions dedicated to not only language processing but executive function as well (Crinion et al., 2006; Kim, Relkin, Lee & Hirsch, 1997; Krizman et al., 2012). On the other hand, bilingual executive function advantages have been argued to be restricted to specific circumstances (Paap

& Greenberg, 2013). Recent meta-analytic reviews reported no evidence for a cognitive control advantage among bilingual adults across ages after publication bias is corrected (e.g., Lehtonen et al., 2018). This leaves the possibility that bilinguals' enhanced cognitive ability may not facilitate auditory perception.

To further our understanding of bilingual enhanced acoustic sensitivity, cognitive ability, and their interplay, we adopted a three-domain by three-language design, where tone (Mandarin Chinese), non-tone (Dutch) late bilingual and non-tone (Dutch) simultaneous bilingual listeners were tested with lexical tone and musical pitch discrimination tasks as well as tasks measuring executive functions. Specifically, we selected non-musicians as participants to limit the influence of music training. Lexical tone discrimination was examined with an AX paradigm using acoustically salient and non-salient contrasts (Liu & Kager, 2014). Music perception was measured by the standard battery, the Montreal Battery on Evaluation of Amusia (MBEA) (Peretz et al., 2003). Executive function and specifically inhibition and switching abilities were measured using Stroop, Trail Making, and Digit Symbol Substitution tests. We ask whether listeners' lexical tone perception, musical pitch discrimination and cognitive abilities differ as a function of different language experiences. We hypothesize higher accuracy in lexical tone and music processing for Chinese and DSB than for Dutch listeners. We further hypothesize that simultaneous bilinguals perform better in executive function tasks than their sequential bilingual peers. As for the interplay, we hypothesize that pitch perception is more unified across linguistic and musical domains for non-tone language listeners. For Chinese listeners, a potential correlation in pitch perception will favour the unified hypothesis on cross-domain perception, and lack of such correlation will mark the impact of phonological categories established through previous native language experience.

Methods

Participants

Twenty-four Dutch (μ : 22.79 years, 6 male), 24 Dutch simultaneous bilingual (DSB, μ : 23.71 years, 9 male) and 24 Mandarin Chinese (μ : 27.50 years, 7 male) speakers were recruited in Utrecht area to participate in the experiment. All participants reported normal hearing sensitivity and no previous history of neurological or psychiatric illnesses. Dutch and Chinese participants had comparable years of exposure to English as late learners (μ_{Dutch} : 14.38, $\mu_{Chinese}$: 16.33, $p = .178$). By contrast, DSB speakers had consistent exposure to another native language since birth (see Appendix I), and used both of their native languages on a regular basis in daily

activities. None of the Dutch participants had experience in a tone or pitch-accent language. The three groups of participants did not differ in overall formal musical training in years ($\mu_{Dutch} : 2.88, \mu_{Chinese} : 2.54, \mu_{DSB} : 2.83, p = .948$). None of the participants was a musician, nor had they received music instruction, training or practice adjacent to the time of the experiment. Participants were tested on 3 sets of experiments in order to measure their linguistic pitch sensitivity, executive function and musical sensitivity.

Experiment 1. Linguistic pitch perception

Stimuli

A female native speaker of Mandarin Chinese recorded tokens of monosyllable /ta/ in isolation, four carrying the high-level (T1) and four the high-falling (T4) using the computer program Audacity via a Genelec 1029A Active Speaker system in a sound-proof booth at the phonetics lab of Utrecht University. The four naturally-produced T1-T4 pairs (Figure 1, contrast A = T1-T4) were further manipulated in PRAAT (Boersma & Weenink, 2009). First, stimuli intensity (65dB) and duration (376 ms) were controlled. Second, four interpolation points along each pitch contour (at 0%, 33%, 67% and 100%) were introduced. Third, the F_0 values occurring at 3/8 and 3/4 of the pitch distance of the original T1 & T4 tokens were calculated at these interpolation points (Appendix II). Linking these points, eight new pitch contours / four new pairs were generated (Figure 1, contrast B = contracted T1-T4). Contrast B thus shrunk the perceptual distance between the two tokens of contrast A. In other words, the acoustic salience of contrast B was weakened solely by the manipulation of F_0 , making them more “difficult” to perceive than contrast A which was perceptually “easy” (Liu & Kager, 2014, 2016). Five native speakers of Mandarin Chinese listened to the stimuli in the environmental settings and ensured that they sounded natural. Moreover, the eight pairs (four contrast A and four B) were presented in two orders (e.g., T1-T4, T4-T1) each, resulting in a total of 16 pairs to be used in the AX discrimination task.

Procedure

To reduce the cognitive load on auditory memory for non-tone language listeners who are unfamiliar with non-native tones, we selected AX over AXB or other discrimination tasks (Gerrits & Schouten, 2004). Participants finished a self-paced AX discrimination task in a sound-proof booth at the phonetic lab of Utrecht University. Participants heard one pair of stimuli per trial, and were required

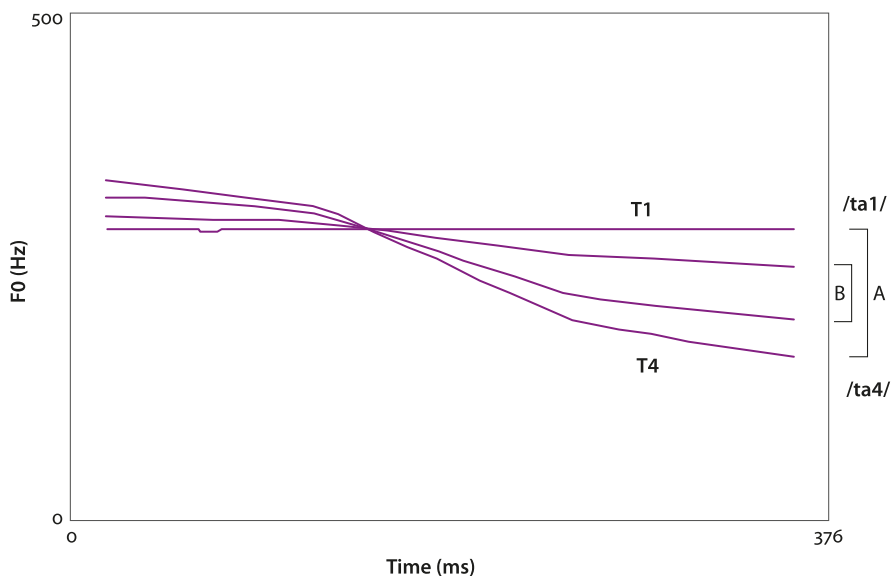


Figure 1. Pitch contours of an illustration of contrast A (T1–T4) and contrast B (contracted T1–T4). The two contrasts differ in pitch distance and subsequently acoustic salience and perceptual difficulty (A more salient than B). (Source from Liu & Kager, 2014)

to respond as accurately and quickly as possible by mouse clicking one of the two buttons on the screen, labelled “same” and “different”, upon hearing each pair. Participants selected “same” when they could not discriminate the paired tokens. If they were able to distinguish the tokens, they chose “different” instead. After each click, the next trial was presented. The inter-stimulus interval between tokens in each pair was set at 500 ms. We did not choose a longer ISI as we did not expect different results if the ISI was longer (Wayland & Guion, 2004). In addition to the different pairs, participants’ responses to the same pairs (identical tokens as filler trials, e.g., T1–T1) were also recorded but results were excluded from analyses. Prior to the task, two auditory examples, T1–T1 and T1–T4, were played as practice trials.

Results

A generalized linear mixed-effects (*glmer*) model was conducted with (binomial) accuracy as dependent variable, language background (3-level, Dutch, DSB, Chinese) and contrast (2-level, contrast A, contrast B) as fixed factors, and participant as the random factor. Both language background (estimate = 4.462, SE = 0.626, $z = 7.133$, $df = 2$, $p < .001$) and contrast (estimate = 4.877, SE = 0.625, $z = 7.804$, $df = 1$,

$p < .001$) were significant predictors of the model. Posthoc pairwise comparisons showed no difference across Dutch, DSB and Chinese participants when discriminating contrast A (estimates < 0.574 , SEs $< .875$, z s < 0.736 , p s $> .742$). When perceiving contrast B (Figure 2), Dutch speakers' performance was significantly lower than their Chinese (estimate = 0.862 , SE = 0.333 , $z = 2.585$, $p = .026$) peers whereas DSB listeners stood in the middle with no significant difference from the other two groups (estimates < 0.573 , SEs $< .335$, z s < 1.714 , p s $> .200$). The means and standard deviations (SD) of the accuracy percentile across groups are listed in Table 1. Note that participants' accuracy scores were not compared against the chance level as they were instructed to select "same" (the inaccurate answer in this case) instead of randomly choosing between the two answers when they could not discriminate the contrast.

Table 1. The means and SD of the accuracy percentile across groups

	Group	Number	Mean Accuracy	SD
Contrast A	Dutch	24	98.44%	4.22%
	DSB	24	96.88%	6.65%
	Chinese	24	98.44%	5.61%
Contrast B	Dutch	24	41.15%	24.02%
	DSB	24	53.65%	29.60%
	Chinese	24	59.90%	18.05%

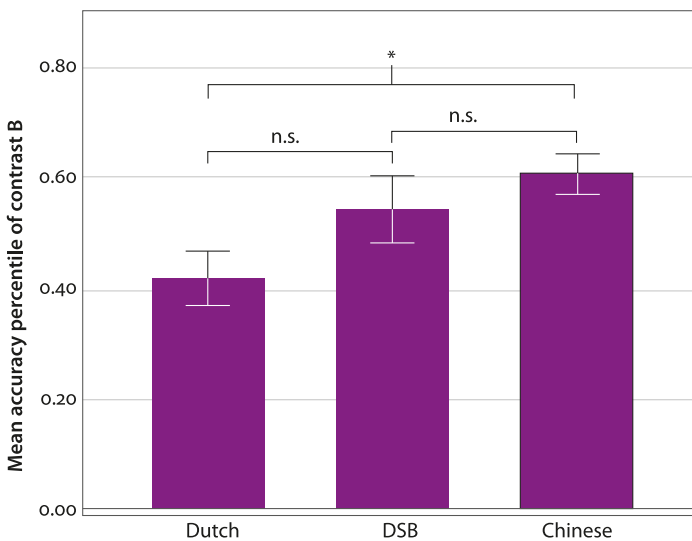


Figure 2. Mean accuracy percentile of contrast B in the discrimination task (Error bar: $\pm 1SE$)

Discussion

Participants from all language backgrounds exhibited ceiling performance with contrast A, in line with previous studies across ages from infancy (Liu & Kager, 2014, 2016) to adulthood (Huang & Johnson, 2011). This is probably due to the intrinsic acoustic properties of the contrast, frequently referred to as acoustic salience (Yeung, Chen & Werker, 2013; Liu & Kager, 2015; Liu, Chen & Kager, 2017). When facing contrast B, a non-native pitch contrast with lower acoustic salience than contrast A, Chinese participants outperformed their Dutch peers and DSB listeners stood in the middle in between the other two groups, reflecting differences in perceptual sensitivity.

Experiment 2. Music perception

Instrument

The Montreal Battery of Evaluation of Amusia (MBEA, Peretz et al., 2003) was employed to measure participants' music perception. The task consisted of six tests: contour, scale, interval, rhythm, metrics/meter and memory. The first three tests targeted listeners' melodic/pitch organization of music, where listeners were required to identify the melodic contour and the tonal functions of the successive pitches. The contour test measured listeners' sensitivity to pitch direction, the scale test, their tonal knowledge, and the interval test, their sensitivity to pitch height. This was followed by two tests assessing listeners' temporal organization of music, representing the metric organization and rhythmic structure of the successive durations. The rhythm test examined listeners' sensitivity to tendencies of temporal proximity without regard to periodicity, whereas the meter test measured their ability to extract underlying temporal regularities or beats. In other words, the three melodic components measured listeners' discrimination of changes in the auditory musical input and the two temporal contents examined their sensitivity to when a musical event occurred. The last test, memory, assessed listeners' music memory capacity.

Each test contained 30–31 questions. In the first five tests of the MBEA, participants were required to make binary (same/different) judgments on two musical phrases of which the second may or may not involve a musical violation. The last test was a memory test in which participants made binary (yes/no) judgments whether a musical phrase was played in the preceding tests.

Procedure

Participants finished the MBEA in a sound-proof booth at the phonetic lab of Utrecht University. Optional breaks were provided in-between every two tests. Most of the participants completed the test in 40 to 60 minutes.

Results

A Linear Mixed-effects Model (LMM) analysis was conducted with participants as the random factor, test (6-level repeated, Contour, Scale, Interval, Rhythm, Meter, Memory), language background (3-level, Dutch, DSB, Chinese), and their interaction as fixed factors, and mean accuracy as the dependent variable. Participants' background ($F(2,72)=3.342, p=.041$), test ($F(5,360)=19.631, p<.001$) and the interaction of language background and test ($F(10,360)=2.408, p=.009$) all significantly predicted performances. Bonferroni corrected pairwise comparisons showed no difference across Dutch, DSB and Chinese participants in Rhythm, Meter or Memory tests, or between DSB and Chinese participants in the other three tests ($ps>.144$). On the other hand, Dutch speakers' scores were significantly or marginally lower than the other two groups in Contour ($p_{\text{DSB}}=.063, p_{\text{Chinese}}=.055$), Scale ($p_{\text{DSB}}=.022, p_{\text{Chinese}}=.010$), and Interval ($p_{\text{DSB}}=.071, p_{\text{Chinese}}=.007$) tests.

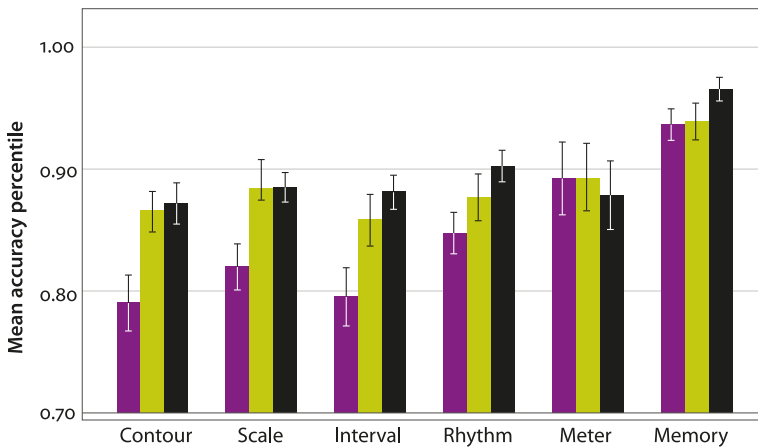


Figure 3. Mean accuracy percentile of the six tests in the MBEA test for Dutch (purple), DSB (green) and Chinese (black) listeners (Error bar: $\pm 1\text{SE}$)

Discussion

The Chinese and DSB groups showed equal performances in the Contour, the Scale, and the Interval tasks, outperforming Dutch participants. Crucially, these three tasks involve pitch perception and discrimination. Tone language (e.g., Chinese) listeners have been shown to perform better at music perception than non-tone (e.g., Dutch) language listeners (Chen et al., 2016). A cross-domain transfer/facilitation effect from linguistic to musical pitch existed for tone language listeners in which word-level pitch differentiate meanings. The finding that DSB listeners outperformed their Dutch peers in pitch-related musical tests was novel, in line with the recent study showing a bilingual advantage in infancy when perceiving violin tunes (Liu & Kager, 2017b). We hypothesize that bilingual advantages within (Hopp, Vogelbacher, Kieseier & Thoma, 2019) and beyond linguistic fields (Liu & Weidemann, 2017) extend to their sensitivity to acoustic cues.

Experiment 3. Executive function

Instrument and procedure

Six tests were conducted targeting participants' executive function and specifically inhibition and switching abilities: Stroop tests (colour, English, native language); Trail Making tests (number, number-letter) and Digit Symbol Substitution test. Proper instructions were given to the participants prior to each test with details listed below.

Stroop tests

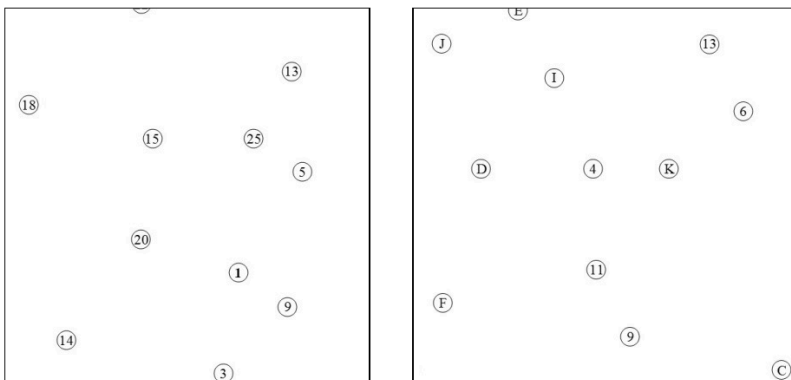
Widely used to test the executive control and conflict resolution differences across ages (McDowd & Shaw, 2000; Bialystok, 2009), the Stroop test effects demonstrate reaction time interference facing incongruent information between the name of a colour and the colour it is printed on (Stroop, 1935). A mismatch between the two (i.e., the word "blue" printed in red ink instead of blue ink) results in longer reaction time and more errors.

Three variations of the Stroop tests were adopted in the current study. In all variations, three types of ink colours (red, green, and blue) were involved. In the Stroop *Colour* test, participants were required to name the ink-colour on "XXXXX" marker with no word meaning. In the Stroop *English* test, participants named the ink-colour of the words ("red", "green", and "blue") while ignoring the word meaning. English was chosen since this single orthography was available for all groups of listeners. The settings of the Stroop *Native* test was similar to the

Stroop English test except that the printed words become Dutch (“rood”, “groen”, and “blauw”) for Dutch and DSB speakers and Chinese characters (“红”, “绿”, and “蓝”) for the Chinese group. The number of correctly named colours within 45 seconds was measured for each test as the dependent variable. The higher the number of correct responses a participant had in the Stroop tests, the better executive function she was assumed to have.

Trail making tests

The Trail Making tests have been used in executive function assessments (Arnett & Labovitz, 1995; Shunk, Davis & Dean, 2006) measuring participants’ flexibility of thinking on visual-motor sequencing. In each test, participants were required to connect 25 consecutive targets in sequential order on a sheet of paper. In the Number Trail Making test, the targets were all numbers from 1 to 25 (Figure 3). In the Number-letter Trail Making test, the order was instructed as switching between numbers and letters (1-A-2-B-.....-L-13; Figure 4). Participants were instructed to complete the test as fast as possible while maintaining accuracy. Errors were corrected immediately before moving on to the next target and were thus reflected in the completion time. While the Number Trail Making test examined participants’ cognitive processing speed, the Number-letter Trail Making test targeted executive functioning (Tombaugh, 2004; Bowie & Harvey, 2006). The lower the completion time, the faster her response was, and the better executive function she had. Participants’ times (in seconds) taken to complete the two tests were used as the dependent variables.



Figures 4 & 5 Example illustrations of the number Trail Making test (left) and the Number-letter Trail Making Test (right)

Digit Symbol Substitution Test

The Digit Symbol Substitution test has been adopted in the Wechsler Adult Intelligence Scale (Kaufman, Flanagan, Alfonso & Mascolo, 2006) and used to test participants' neuropsychological conditions (Lezak, Howieson & Loring, 2004). The test used in the current study consisted of nine digit-symbol pairs (Figure 5) followed by a list of digits. Participants were required to write down the corresponding symbols as fast as possible within 90 seconds. The number of correct symbols written was measured as the dependent variable. The higher the number of symbols the participant wrote down, the better her response was, and the better executive function she had.



Figure 6. Nine digit-symbol pairs used in the Digit Symbol Substitution test

Results

Stroop tests

An LMM analysis was conducted with participants as the random factor, task (3-level repeated, Colour, English, Native), language background (3-level, Dutch, DSB, Chinese), and their interaction as fixed factors, and number of correct responses as the dependent variable. Participants' background did not predict participants' scores ($F(2,72)=0.155$, $p=.856$) while Stroop ($F(2,144)=180.451$, $p<.001$) and the interaction of language background and the task ($F(4,144)=16.152$, $p<.001$) significantly predicted participants' scores. Bonferroni corrected pairwise comparisons showed no difference across Dutch, DSB and Chinese participants in the Colour task ($ps>.05$), a marginal advantage for Chinese participants in the English task over DSB ($p=.054$), and a weak performance was observed for Chinese participants compared to the other two groups ($ps<.048$) in the Native task (Figure 6). No difference was observed between Dutch and DSB groups across tasks ($ps>.099$). Cross-test correlation showed that participants' results of each test were always correlated with one another across each language condition ($rs>.667$, $ps<.001$, $N=24$).

Trail making tests

An LMM analysis was conducted with participants as the random factor, task (2-level repeated, Number, Number-letter), language background (3-level, Dutch,

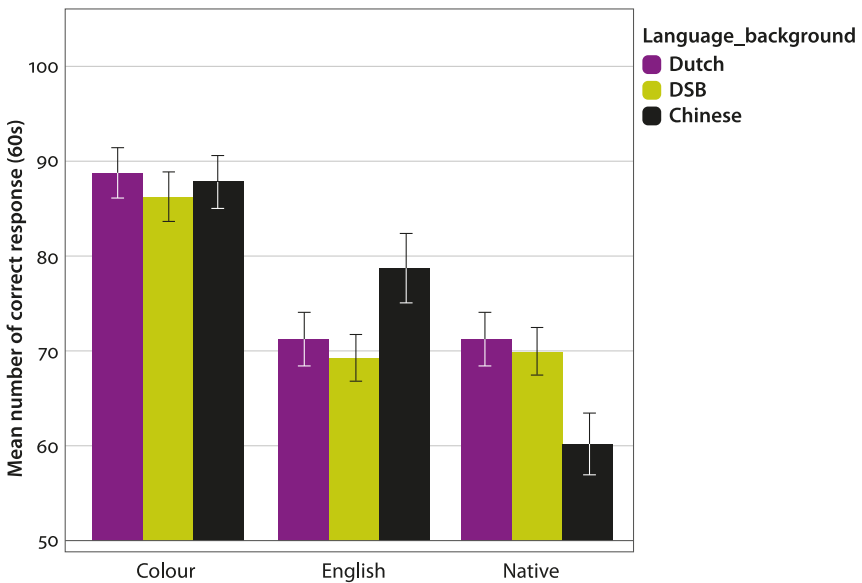


Figure 7. Mean number of correct responses within 60 seconds in the three Stroop tests (Error bar: ± 1 SE). The longer the bar, the better the performance

DSB, Chinese), and their interaction as fixed factors, and completion time as the dependent variable. Participants' background did not predict performance ($F(2, 72) = 1.714, p = .187$) while task ($F(1, 72) = 94.320, p < .001$) and the interaction of language background and the task ($F(2, 72) = 4.663, p = .012$) significantly predicted participants' scores. Bonferroni corrected pairwise comparisons showed no difference across Dutch, DSB and Chinese participants in the Number task ($ps > .05$), whereas a weak performance was observed for Chinese participants compared to the other two groups ($ps < .044$) in the Number-letter task (Figure 7). Cross-test correlation revealed that participants' results of the two tests were correlated in each group ($rs > .549, N = 24, ps < .006$).

Digit symbol substitution test

An LMM analysis was conducted with participants as the random factor, language background (3-level, Dutch, DSB, Chinese) as the fixed factor, and number of correct responses as the dependent variable. Although participants' background predicted participants' scores ($F(2, 72) = 3.139, p = .049$), Bonferroni corrected pairwise comparisons showed no difference across language backgrounds (Dutch-DSB: $p > .999$, Dutch-Chinese: $p = .093$, DSB-Chinese: $p = .108$, Figure 8). Results indicated a trend for Chinese over Dutch and DSB listeners, although the differences were small.

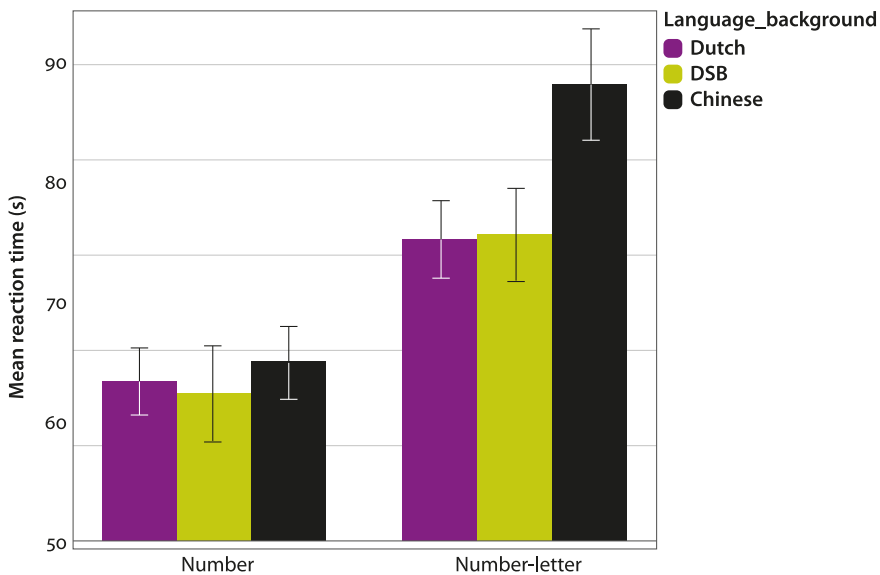


Figure 8. Mean completion time in seconds for the two Trail Making tests (Error bar: $\pm 1SE$). The shorter the bar, the better the performance

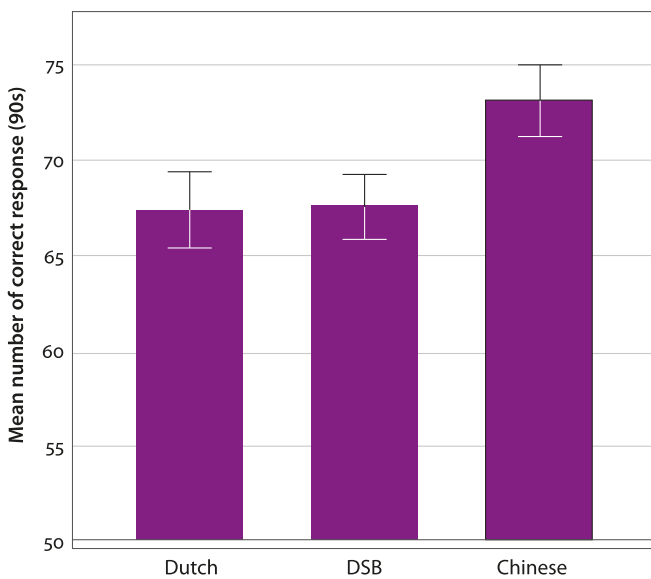


Figure 9. Mean number of correct responses within 90 seconds in the Digit Symbol Substitution Test (Error bar: $\pm 1SE$)

Discussion

Three sets of experiments targeting participant executive function presented two main trends. First, there was no robust differences in the majority of tests across three groups of participants. Second, Chinese participants demonstrated a disadvantage in the Stroop Native test and the Trail-marking number-letter test. We discuss these two trends separately.

Our first trend indicates no evidence supporting a bilingual advantage in executive function. Although participants were young adults, data did not seem to suggest they achieved their ceiling performance. Relating to the debate regarding cognitive benefits of bilingualism (Bialystok & DePape, 2009; De Bruin, Treccani & Della Sala, 2014; Von Bastian, Souza & Gade, 2016), our results conform to studies and meta-analyses on executive function-related tasks all reporting no bilingual advantage in children (Dick et al., 2019; Duñabeitia et al., 2014), young adults (Costa, Hernández, Costa-Faidella & Sebastián-Gallés, 2009; Donnelly et al., 2015), old adults (Lukasik et al., 2018) or across ages (Lehtonen et al., 2018; Paap et al., 2015) and differ from others reporting a bilingual advantage in cognitive abilities (Bak, Nissan, Allerhand & Deary, 2014; Vega-Mendoza, West, Sorace & Bak, 2015). However, our data do not reject the possibility that the lack of evidence is related to participants' age, as young adult samples have consistently not shown the monolingual versus bilingual differences in executive function (Antoniou, 2019; Bialystok, 2017; Valian, 2015).

Secondly, the performances of Chinese participants may reveal some degree of cognitive interference of language processing. They were affected when words were presented in their native language and in English. Such interference does not appear to be equal in magnitude across participants: Dutch and DSB participants seemed less affected than their Chinese peers. The overall finding may be attributed to different cognitive demands in the processing of different orthographic systems (Wang, Perfetti & Liu, 2005): the processing of logograms being qualitatively different from words in the alphabetic script. That is, the processing of Chinese characters may be more "automatic" and therefore requires more cognitive effort to inhibit. Alternatively, processing logographs may be more effortful and therefore require higher cognitive resources, which would explain why Mandarin speakers show low levels of performance in the native version. In another test, Number-letter Trail Making, Chinese participants required a longer time to process the number-letter alteration compared to their Dutch and DSB peers, whereas everyone performed the same in the number-only condition. While the familiarity with the order of numbers (1-2-3-4) may stay equal across participants, users speaking a non-alphabetic native language may have

additional processing cost to recall the order of letters (A-B-C-D) especially when numbers were inserted.

Cross-task correlations

The correlations across linguistic, cognitive and musical pitch perception tasks were calculated. No correlation was observed between participants' overall executive function performances and their linguistic and musical perception ($ps > .05$). Neither was there significant correlation between contrast A and music task results ($ps > .05$) probably due to participants' ceiling performance in the former. Nevertheless, a strong correlation was observed between contrast B and music task results, $rs > .276$, $N = 72$, $ps < .020$ (Table 2). Splitting the data by language background and looking into the relationship between the linguistic contrast and individual musical tasks (Table 3), Dutch participants showed the highest number of significant correlation in music tasks (Contour, Scale, Interval and Memory, $ps < .018$), followed by DSB (Interval and Memory, $p < .022$) and then Chinese participants (Rhythm, $p = .018$). Note that Contour, Scale, and Interval tasks targeted one's sensitivity to musical pitch. The interplay between linguistic and musical pitch perception appeared to be the strongest among Dutch and the weakest among Chinese participants. Similar to previous findings, DSB listeners stood in the middle between the other two groups with a moderate correlation between speech and music perception. The lack of correlation between speech and music perception among Chinese and DSB listeners is likely to be due to participants' overall good performances (Figures 10–12). Compared to Dutch and DSB listeners, the correlation between Chinese listeners' linguistic pitch and musical rhythm perception may stem from the fundamental rhythmic difference of their native languages. Chinese is a syllable-timed language whereas Dutch is stress-timed (Abercrombie, 1967; Mok, 2009). Twenty participants (except for two French, one Spanish and one Tamil) from the DSB group speak two stress-timed languages.

General discussion

This study examined whether listeners' lexical tone perception, musical pitch discrimination and cognitive abilities differ as a function of different language experiences, and explored the interplay between listeners' performances. Results are expected to lead to our advanced understanding of whether bilingual advantages in speech and music perception observed in infancy extend to adulthood, whether the lexical tone-musical pitch correlation observed among non-tone

Table 2. Correlation between multiple tasks performed by Dutch listeners

	ContrastB	Contour	Scale	Interval	Rhythm	Meter	Memory	Stroop	Trial	DSS
ContrastA	$r=.19$	$r=.06$	$r=.50^*$	$r=.29$	$r=-.06$	$r=.24$	$r=-.18$	$r=-.03$	$r=-.11$	$r=.11$
ContrastB		$r=.48^*$	$r=.49^*$	$r=.58^{**}$	$r=.31$	$r=.39$	$r=.51^*$	$r=-.27$	$r=.26$	$r=-.35$
Contour			$r=.60^{**}$	$r=.73^{**}$	$r=.67^{**}$	$r=.25$	$r=.49^*$	$r=.07$	$r=-.10$	$r=-.30$
Scale				$r=.61^{**}$	$r=.57^{**}$	$r=.60^{**}$	$r=.29$	$r=-.25$	$r=.05$	$r=-.01$
Interval					$r=.73^{**}$	$r=.08$	$r=.33$	$r=.07$	$r=.19$	$r=-.35$
Rhythm						$r=.12$	$r=.42^*$	$r=-.11$	$r=.07$	$r=-.13$
Meter							$r=.45^*$	$r=-.39$	$r=.01$	$r=.26$
Memory								$r=-.33$	$r=-.02$	$r=-.12$
Stroop									$r=-.29$	$r=.05$
Trial										$r=-.63^{**}$

* Significance at .05 or lower (two-tailed) ** Significance at .01 or lower (two-tailed)

Table 3. Correlation between multiple tasks performed by DSB listeners

	ContrastB	Contour	Scale	Interval	Rhythm	Meter	Memory	Stroop	Trial	DSS
ContrastA	$r=.54^{**}$	$r=.28$	$r=.19$	$r=.25$	$r=.15$	$r=.24$	$r=.37$	$r=-.18$	$r=-.28$	$r=-.01$
ContrastB		$r=.39$	$r=.18$	$r=.49^*$	$r=.34$	$r=.38$	$r=.47^*$	$r=-.24$	$r=.17$	$r=-.24$
Contour			$r=.82^{**}$	$r=.74^{**}$	$r=.47^*$	$r=.47^*$	$r=.64^{**}$	$r=-.36$	$r=.11$	$r=-.20$
Scale				$r=.63^{**}$	$r=.26$	$r=.30$	$r=.53^{**}$	$r=-.18$	$r=.15$	$r=-.20$
Interval					$r=.70^{**}$	$r=.47^*$	$r=.69^{**}$	$r=-.20$	$r=.16$	$r=-.04$
Rhythm						$r=.63^{**}$	$r=.55^{**}$	$r=.09$	$r=.21$	$r=-.28$
Meter							$r=.42^*$	$r=-.04$	$r=-.01$	$r=-.30$
Memory								$r=-.27$	$r=.16$	$r=-.15$
Stroop									$r=-.16$	$r=-.04$
Trial										$r=-.57^{**}$

* Significance at .05 or lower (two-tailed) ** Significance at .01 or lower (two-tailed)

language listeners (Chen, Liu & Kager, 2015) may extend to simultaneous bilinguals, and to what extent these listeners' perception may be influenced by cognitive factors.

This study examined linguistic pitch, music perception and executive function of listeners from three types of language backgrounds. Results demonstrated that: First, while all participants were able to discriminate between salient linguistic pitch contours, Chinese and outperformed their Dutch peers when the acoustic salience of the contrast was diminished, and DSB listeners' performance

Table 4. Correlation between multiple tasks performed by Chinese listeners

	ContrastB	Contour	Scale	Interval	Rhythm	Meter	Memory	Stroop	Trial	DSS
ContrastA	$r=.50^*$	$r=.07$	$r=-.15$	$r=.14$	$r=.55^{**}$	$r=.12$	$r=.14$	$r=.38$	$r=-.23$	$r=.04$
ContrastB		$r=.19$	$r=-.12$	$r=.05$	$r=.48^*$	$r=.08$	$r=.04$	$r=.11$	$r=-.10$	$r=-.05$
Contour			$r=-.71^{**}$	$r=-.75^{**}$	$r=.50^*$	$r=.35$	$r=.33$	$r=.19$	$r=.06$	$r=-.01$
Scale				$r=.82^{**}$	$r=.17$	$r=.24$	$r=.66^{**}$	$r=.11$	$r=.11$	$r=.14$
Interval					$r=.32$	$r=.37$	$r=.71^{**}$	$r=.23$	$r=-.19$	$r=.18$
Rhythm						$r=.38$	$r=.13$	$r=.26$	$r=-.06$	$r=.02$
Meter							$r=.31$	$r=.19$	$r=-.17$	$r=-.29$
Memory								$r=.31$	$r=-.05$	$r=.20$
Stroop									$r=.01$	$r=-.07$
Trial										$r=-.35$

* Significance at .05 or lower (two-tailed) ** Significance at .01 or lower (two-tailed)

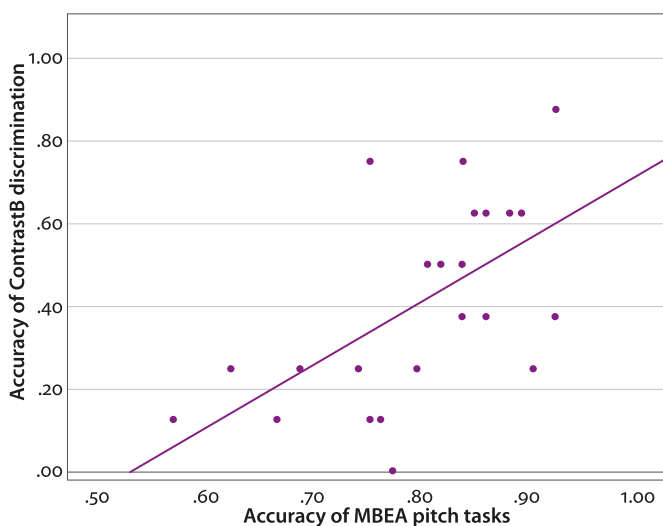


Figure 10. Correlation between the lexical tone (ContrastB) discrimination tasks and the MBEA pitch tasks for the Dutch listeners. Linear $R^2=0.36$

stood in the middle. Second, Chinese and DSB listeners did not differ in their MBEA performance and outperformed Dutch listeners once again in pitch-related musical tasks. Third, there was no evidence of overall executive function differences across participants, nor did their executive function interact with pitch perception. Last but not least, the correlation between participants' linguistic and musical pitch sensitivity was the most evident among Dutch listeners,

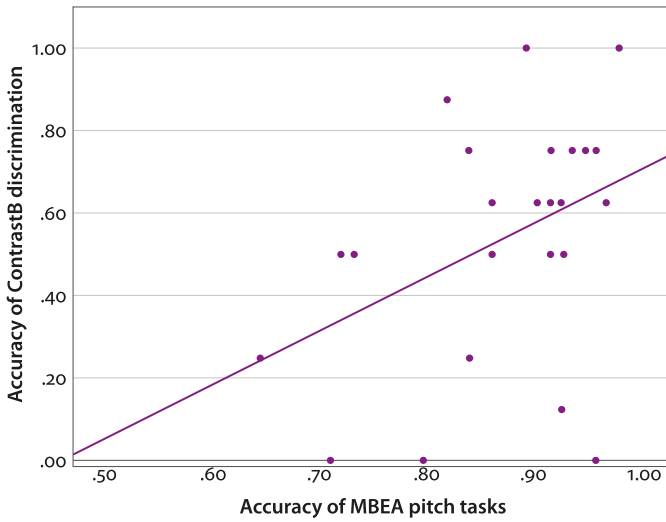


Figure 11. Correlation between the lexical tone (ContrastB) discrimination tasks and the MBEA pitch tasks for the DSB listeners. Linear $R^2 = 0.16$

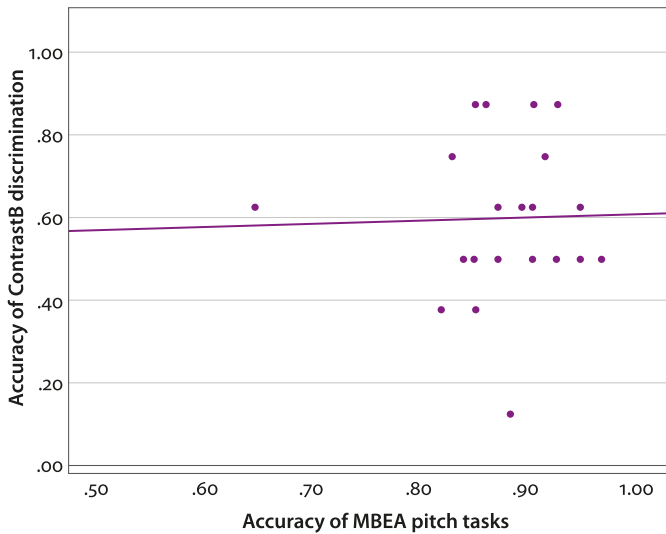


Figure 12. Correlation between the lexical tone (ContrastB) discrimination tasks and the MBEA pitch tasks for the Chinese listeners. Linear $R^2 = 0.01$

followed by DSB and Chinese listeners who demonstrated limited and no cross-domain interplay, respectively.

In line with previous findings illustrating the positive effect of tone language exposure on lexical and musical pitch perception (Bidelman et al., 2011a; Pfordresher & Brown, 2009), Chinese participants' linguistic pitch perceptual advantage is as expected, because it involved a native tone contrast. Since the differences between DSB and Dutch listeners did not reach significance, the limited enhancement trend in the DSB group needs to be interpreted with caution. Previous studies have demonstrated that just like music experience (Thompson, Schellenberg & Husain, 2004), life-long exposure to a bilingual environment may improve the subcortical encoding of sounds including those beyond native languages (Krizman et al., 2012). In fact, bilinguals' enhanced neural and acoustic sensitivity to non-native speech contrasts is present in the first year of life (Antoniou et al., 2015; Liu & Kager, 2016; Petitto et al., 2012; Ressel et al., 2012). It is not surprising that such a benefit can be carried over to adulthood given listeners' cumulative language exposure. Although the DSB listeners in the current experiment had no experience in a specific tone or pitch-accent language, they nevertheless, typically experience more prosodic variation (e.g., intonation) and develop refined sound categories from two languages (Kuhl et al., 2008; Liu & Kager, 2015). This richer prosodic experience may subsequently facilitate their perception of other suprasegmental features. Following the heightened acoustic sensitivity hypothesis (Liu, 2014; Liu & Kager, 2016), we argue that simultaneous bilingual experience sharpens listeners' overall sensitivity to speech sounds. More importantly, such sharpening is not restricted to perceiving the phonetic details which bilinguals are specifically exposed in their native languages, but extending to inexperienced acoustic features and other domains, evidenced by the current results showing non-tone language bilinguals' enhanced performances in lexical tones and musical pitch. Infant research has reported bilingual advantages in speech perception involving linguistic pitch (Hay, Graf Estes, Wang & Saffran, 2015; Singh, 2017; Singh, Fu, Tay & Golinkoff, 2017). Whether or not listeners assimilate non-native linguistic pitch to native intonation is an issue of debate. Some studies find evidence of perceptual assimilation (Chen & Kager, 2016; Singh & Chee, 2016) when testing (adult and child) non-tone language listeners' perception of tonal contrasts that resemble native intonation (e.g., rising vs. falling tone, resembling interrogative and narrative intonation). Others, however, report results showing no intonation facilitation effect as such (Tsao, 2017). Another related issue is whether the observed bilingual advantage can be more specifically attributed, for instance, to similarities/differences in speech prosody (word stress, intonation, rhythmicity) between the two native languages. Though scarce, previous studies have reported enhanced sensitivity in stress and rhythm processing for bilinguals from infancy (Molnar et al., 2014) to adulthood (Abboub, Bijeljac-Babic, Serres & Nazzi, 2015; Molnar, Carreiras & Gervain, 2016; also see

Roncaglia-Denissen, Roor, Chen & Sadakata, 2016 for similar enhancement effects in L2A). We leave these possibilities open for future studies.

In addition, differences in music perception have been observed between late and simultaneous bilingual Dutch listeners, the latter being more accurate in pitch-related tasks than the former. As enhanced music sensitivity surfaces in bilingual infants of 9 months (Liu & Kager, 2017b), bilingual enhanced acoustic sensitivity may be an age- and domain-general effect. The magnitude of bilingual's heightened acoustic sensitivity may, however, depend on variables such as their specific auditory experience and input acoustic properties. The reason why such advantages surface in musical pitch perception could be that pitch is a relatively salient feature (Yeung et al., 2013), as very young infants show sensitivity to word-level pitch variations (Nazzi et al., 1998) and melodic consonance (Trainor, Tsang & Cheung, 2002). It has been demonstrated that both linguistic and musical expertise benefit music perception: Chinese listeners and non-tone language musicians have stronger brainstem representation of the defining pitches of musical sequences than non-musicians (Bidelman, Gandour & Krishna, 2011b). Our findings suggest that a bilingual environment may enhance listeners' general auditory perception, possibly resulting in the enhanced subcortical encoding of incoming acoustics (Krizman et al., 2012).

Our initial hypothesis that the nature of bilingual enhanced sensitivity can be explained from cognitive perspectives was not supported. The current study did not find any executive function advantage or correlation between language/music perception and executive function among DSB listeners. In addition to the heated debate on the magnitude of bilingual cognitive advantage (e.g., Bialystok, 2017; Paap et al., 2015), the lack of correlation may also be attributed to the specific tasks conducted in the current study: Although the Stroop task requires the production of three colour words, none of the current executive function tasks entails acoustic perception. Participants' executive function ability may be more relevant to tasks involving high-level cognitive function and less relevant to the auditory perceptual tasks examining lower-level acoustic mechanisms. We leave the executive function debate and its function open for future research.

Finally, the strength of the correlation between speech and music perception differed significantly across listeners from various backgrounds in this study. Dutch listeners' speech and music perception were highly linked, replicating the previous study (Chen et al., 2016). Such correlations indicate the involvement of shared or parallel mechanisms in listeners' linguistic pitch and music perception (Hallé, Chang & Best, 2004), suggesting cross-domain perceptual transfer (Bialystok & Depape, 2009; Bidelman et al., 2011a; Moreno, 2009). Comparatively speaking, DSB listeners' strength of correlation was much less robust (well associated) than that in Dutch listeners, and lexical tone and music perception were

hardly correlated in Chinese listeners, indicating the influence of native phonemic categories. Nevertheless, the latter did demonstrate the highest sensitivity in linguistic pitch and music among the three groups, illustrating cross-domain benefits stemming from native tone language exposure. In line with the current finding, Cooper and Wang (2012) found that English (non-tone language) but not Thai (tone language) listeners' tone identification and musical aptitude scores were significantly correlated with their successful learning of novel tonal words in Cantonese. It could be that once the lexical tones are fully acquired, they are/become isolated from other non-lexical pitch variations, leading to no correlation between lexical tone and musical pitch processing. To conclude, the overall study contributes to the existing literature by demonstrating a bilingual advantage in pitch perception across domains of language and music.

Funding

Research funded by Utrecht University to Liquan Liu and by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie (798658) to Liquan Liu for writing.

References

- Abercombie, D. (1967). *Elements of general phonetics*. Aldine Pub. Company.
- Abboub, N., Bijeljac-Babic, R., Serres, J., & Nazzi, T. (2015). On the importance of being bilingual: Word stress processing in a context of segmental variability. *Journal of Experimental Child Psychology*, 132, 111–120. <https://doi.org/10.1016/j.jecp.2014.12.004>
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child neuropsychology*, 8(2), 71–82. <https://doi.org/10.1076/chin.8.2.71.8724>
- Antoniou, M. (2019). The advantages of bilingualism debate. *Annual Review of Linguistics*, 5, 395–415. <https://doi.org/10.1146/annurev-linguistics-011718-011820>
- Antoniou, M., Liang, E., Ettliger, M., & Wong, P.C. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, 18(4), 683–695. <https://doi.org/10.1017/S1366728914000777>
- Arnett, J.A., & Labovitz, S.S. (1995). Effect of physical layout in performance of the Trail Making Test. *Psychological Assessment*, 7(2), 220. <https://doi.org/10.1037/1040-3590.7.2.220>
- Bak, T.H., Nissan, J.J., Allerhand, M.M., & Deary, I.J. (2014). Does bilingualism influence cognitive aging?. *Annals of Neurology*, 75(6), 959–963. <https://doi.org/10.1002/ana.24158>
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, 70(3), 636–644. <https://doi.org/10.1111/1467-8624.00046>
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(01), 3–11. <https://doi.org/10.1017/S1366728908003477>
- Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychological bulletin*, 143(3), 233. <https://doi.org/10.1037/bul0000099>

- Bialystok, E., Craik, F.I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging, 19*(2), 290. <https://doi.org/10.1037/0882-7974.19.2.290>
- Bialystok, E., & DePape, A.M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 565.
- Bialystok, E., Martin, M.M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism, 9*(1), 103–119. <https://doi.org/10.1177/13670069050090010701>
- Bidelman, G.M., & Chung, W.L. (2015). Tone-language speakers show hemispheric specialization and differential cortical processing of contour and interval cues for pitch. *Neuroscience, 305*, 384–392. <https://doi.org/10.1016/j.neuroscience.2015.08.010>
- 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*(2), 425–434. <https://doi.org/10.1162/jocn.2009.21362>
- 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), 1–10. <https://doi.org/10.1016/j.bandc.2011.07.006>
- 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*(4), e60676. <https://doi.org/10.1371/journal.pone.0060676>
- Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.05) [Computer program]. Retrieved May 1, 2009.
- Bowie, C.R., & Harvey, P.D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols, 1*(5), 2277–2281. <https://doi.org/10.1038/nprot.2006.390>
- Burnham, D.K., & Brooker, R. (2002). Absolute pitch and lexical tones: Tone perception by non-musician, musician, and absolute pitch non-tonal language speakers. In J. Hansen, & B. Pellom (Eds.), *Seventh International Conference on Spoken Language Processing*.
- Carlson, S.M., & Meltzoff, A.N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*(2), 282–298. <https://doi.org/10.1111/j.1467-7687.2008.00675.x>
- Chen, A., & Kager, R. (2016). Discrimination of lexical tones in the first year of life. *Infant and Child Development, 25*(5), 426–439. <https://doi.org/10.1002/icd.1944>
- Chen, A., Liu, L., & Kager, R. (2015). Cross-linguistic perception of Mandarin tone sandhi. *Language Sciences, 48*, 62–69. <https://doi.org/10.1016/j.langsci.2014.12.002>
- Chen, A., Liu, L., & Kager, R. (2016). Cross-domain correlation in pitch perception, the influence of native language. *Language, Cognition and Neuroscience, 31*(6), 751–760. <https://doi.org/10.1080/23273798.2016.1156715>
- Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *The Journal of the Acoustical Society of America, 131*(6), 4756–4769. <https://doi.org/10.1121/1.4714355>
- Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113*(2), 135–149. <https://doi.org/10.1016/j.cognition.2009.08.001>

- Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T.,... & Usui, K. (2006). Language control in the bilingual brain. *Science*, *312*(5779), 1537–1540. <https://doi.org/10.1126/science.1127761>
- De Bruin, A., Treccani, B., & Della Sala, S. (2014). Cognitive advantage in bilingualism an example of publication bias?. *Psychological Science*, *26*(1), 99–107. <https://doi.org/10.1177/0956797614557866>
- 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*(3), 203–207. <https://doi.org/10.1007/s10339-006-0146-7>
- 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*(1), 46–61. <https://doi.org/10.1080/09541440802708136>
- Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T.,... & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nature human behaviour*, *3*(7), 692–701. <https://doi.org/10.1038/s41562-019-0609-3>
- Donnelly, S., Brooks, P. J., & Homer, B. D. (2015, July). Examining the bilingual advantage on conflict resolution tasks: A meta-analysis. In D. Noelle, & R. Dale (Eds.), *Proceedings of the 37th Annual Conference of The Cognitive Science Society*, Pasadena, CA.
- Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited. *Experimental Psychology*, *61*(3), 234–251. <https://doi.org/10.1027/1618-3169/a000243>
- Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, *36*(2), 268–294. <https://doi.org/10.1016/j.jwocn.2007.06.005>
- Gandour, J., Tong, Y., Wong, D., Talavage, T., Dziedzic, M., Xu, Y., Lie, X., & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage*, *23*(1), 344–357. <https://doi.org/10.1016/j.neuroimage.2004.06.004>
- Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, *66*(3), 363–376. <https://doi.org/10.3758/BF03194885>
- Gioia, G. A., Isquith, P. K., & Guy, S. C. (2001). Assessment of executive functions in children with neurological impairment. In R. J. Simeonsson, & L. Rosenthal (Eds.), *Psychological and Developmental Assessment: Children with Disabilities and Chronic Conditions* (317–356). New York: Guilford Press.
- 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*(3), 395–421. [https://doi.org/10.1016/S0095-4470\(03\)00016-0](https://doi.org/10.1016/S0095-4470(03)00016-0)
- Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, *11*(11), 466–472. <https://doi.org/10.1016/j.tics.2007.08.008>
- 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*(1), 10–22. <https://doi.org/10.1111/cdev.12269>
- Hopp, H., Vogelbacher, M., Kieseffer, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction*, *61*, 99–110. <https://doi.org/10.1016/j.learninstruc.2019.02.001>

- Huang, T., & Johnson, K. (2011). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica*, 67(4), 243–267. <https://doi.org/10.1159/000327392>
- Kaufman, A. S., Flanagan, D. P., Alfonso, V. C., & Mascolo, J. T. (2006). Test review: Wechsler intelligence scale for children, (WISC-IV). *Journal of Psychoeducational Assessment*, 24(3), 278–295. <https://doi.org/10.1177/0734282906288389>
- Kim, K. H., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, 388(6638), 171–174. <https://doi.org/10.1038/40623>
- 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(4), 646–653. <https://doi.org/10.1006/nimg.2000.0738>
- Koelsch, S., Gunter, T. C., Cramon, D. Y. V., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: a cortical “language-network” serves the processing of music. *Neuroimage*, 17(2), 956–966. <https://doi.org/10.1006/nimg.2002.1154>
- Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences*, 109(20), 7877–7881. <https://doi.org/10.1073/pnas.1201575109>
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9(2), F13–F21. <https://doi.org/10.1111/j.1467-7687.2006.00468.x>
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 979–1000. <https://doi.org/10.1098/rstb.2007.2154>
- Lee, C. Y., & Hung, T. H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *The Journal of the Acoustical Society of America*, 124(5), 3235–3248. <https://doi.org/10.1121/1.2990713>
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., De Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological bulletin*, 144(4), 394. <https://doi.org/10.1037/bul0000142>
- Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological evaluation*. New York: Oxford University Press.
- Liu, F., Jiang, C., Thompson, W., 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. (2014). The Effects of Bilingualism on Infant Language Development: The Acquisition of Sounds and Words. PhD dissertation, Utrecht University (ISBN 78-94-6093-129-1).
- Liu, L., Chen, A., & Kager, R. (2017). Tone perception in Mandarin and Dutch adult listeners. *Language and Linguistics*, 18(4), 622–646.
- Liu, L., & Kager, R. W. J. (2014). Perception of tones by infants learning a non-tone language. Perception of tones by infants learning a non-tone language. *Cognition*, 133(2), 185–194. <https://doi.org/10.1016/j.cognition.2014.06.004>
- Liu, L., & Kager, R. W. J. (2015). Understanding phonological acquisition through phonetic perception: The influence of exposure and acoustic salience. *Phonological Studies*, 18, 51–58.

- Liu, L., & Kager, R. W. J. (2016). Perception of a native vowel contrast by Dutch monolingual and bilingual infants: A bilingual perceptual lead. *International Journal of Bilingualism*, 20(3), 335–345. <https://doi.org/10.1177/1367006914566082>
- Liu, L., & Kager, R. W. J. (2017a). Perception of tones by bilingual infants learning non-tone languages. *Bilingualism: Language and Cognition*, 20(3), 561–575. <https://doi.org/10.1017/S1366728916000183>
- Liu, L., & Kager, R. W. J. (2017b). Enhanced music sensitivity in 9-month-old bilingual infants. *Cognitive Processing*, 18(1), 55–56. <https://doi.org/10.1007/s10339-016-0780-7>
- Liu, L., & Weidemann, G. (2017). Is it wise to raise your child bilingually?. *China Language Strategies*, 5(1), Nanjing University Press.
- Lukasik, K. M., Lehtonen, M., Soveri, A., Waris, O., Jylkkä, J., & Laine, M. (2018). Bilingualism and working memory performance: Evidence from a large-scale online study. *PLoS one*, 13(11). <https://doi.org/10.1371/journal.pone.0205916>
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10(3), 241–265. https://doi.org/10.1207/s15327078in1003_3
- McDowd, J. M. & Shaw, R. J. (2000). Attention and aging: A functional perspective.
- McMullen, E. & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception: An Interdisciplinary Journal*, 21(3), 289–311. <https://doi.org/10.1525/mp.2004.21.3.289>
- Mok, P. (2009). On the syllable-timing of Cantonese and Beijing Mandarin. *Chinese Journal of Phonetics*, 2, 148–154.
- Molnar, M., Carreiras, M., & Gervain, J. (2016). Language dominance shapes non-linguistic rhythmic grouping in bilinguals. *Cognition*, 152, 150–159. <https://doi.org/10.1016/j.cognition.2016.03.023>
- Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, 19(3), 326–337. <https://doi.org/10.1111/inf.12041>
- Moreno, S. (2009). Can music influence language and cognition?. *Contemporary Music Review*, 28(3), 329–345. <https://doi.org/10.1080/07494460903404410>
- Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain*, 133(9), 2635–2642. <https://doi.org/10.1093/brain/awq178>
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756.
- Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music* (pp. 463–472). Cambridge, MA: MIT Press.
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232–258. <https://doi.org/10.1016/j.cogpsych.2012.12.002>
- Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, 69, 265–278. <https://doi.org/10.1016/j.cortex.2015.04.014>
- Patel, A. D. (2008). Science & music: talk of the tone. *Nature*, 453(7196), 726. <https://doi.org/10.1038/453726a>

- Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, 11(9), 369–372. <https://doi.org/10.1016/j.tics.2007.08.003>
- Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. *Annals of the New York Academy of Sciences*, 999(1), 58–75. <https://doi.org/10.1196/annals.1284.006>
- Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The “Perceptual Wedge Hypothesis” as the basis for bilingual babies’ phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, 121(2), 130–143. <https://doi.org/10.1016/j.bandl.2011.05.003>
- Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception & Psychophysics*, 71(6), 1385–1398. <https://doi.org/10.3758/APP.71.6.1385>
- Ressel, V., Pallier, C., Ventura-Campos, N., Díaz, B., Roessler, A., Ávila, C., & Sebastián-Gallés, N. (2012). An effect of bilingualism on the auditory cortex. *Journal of Neuroscience*, 32(47), 16597–16601. <https://doi.org/10.1523/JNEUROSCI.1996-12.2012>
- Roncaglia-Denissen, P., Roor, D., Chen, A., & Sadakata, M. (2016). Mastering languages with different rhythmic properties enhances musical rhythm perception. *Frontiers in Human Neuroscience*, 10, 228. <https://doi.org/10.3389/fnhum.2016.00288>
- Shunk, A. W., Davis, A. S., & Dean, R. S. (2006). TEST REVIEW: Dean C. Delis, Edith Kaplan & Joel H. Kramer, Delis Kaplan Executive Function System (D-KEFS), The Psychological Corporation, San Antonio, TX, 2001. \$415.00 (complete kit). *Applied Neuropsychology*, 13(4), 275–327. https://doi.org/10.1207/s15324826an1304_9
- Singh, L. (2017). Bilingual infants demonstrate advantages in learning words in a third language. *Child Development*, 89(4), e397–e413. <https://doi.org/10.1111/cdev.12852>
- Singh, L., & Chee, M. (2016). Rise and fall: effects of tone and intonation on spoken word recognition in early childhood. *Journal of Phonetics*, 55, 109–118. <https://doi.org/10.1016/j.jocn.2015.12.005>
- Singh, L., Fu, C. S., Tay, Z. W., & Golinkoff, R. M. (2017). Novel word learning in bilingual and monolingual infants: Evidence for a bilingual advantage. *Child Development*, 89(3), e183–e198. <https://doi.org/10.1111/cdev.12747>
- Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: A cross-cultural comparison. *Developmental Psychology*, 46(1), 286–292. <https://doi.org/10.1037/a0017555>
- Song, J. H., Skoe, E., Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, 20(10), 1892–1902. <https://doi.org/10.1162/jocn.2008.20131>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, 18(6), 643. <https://doi.org/10.1037/h0054651>
- Tervaniemi, M. A., 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(3), 330–336. <https://doi.org/10.1006/nimg.1999.0405>
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help?. *Emotion*, 4(1), 46. <https://doi.org/10.1037/1528-3542.4.1.46>
- Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of Clinical Neuropsychology*, 19(2), 203–214. [https://doi.org/10.1016/S0887-6177\(03\)00039-8](https://doi.org/10.1016/S0887-6177(03)00039-8)

- Trainor, L. J., Tsang, C. D., & Cheung, V. H. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception: An Interdisciplinary Journal*, 20(2), 187–194. <https://doi.org/10.1525/mp.2002.20.2.187>
- Tsao, F. M. (2017). Perceptual improvement of lexical tones in infants: effects of tone language experience. *Frontiers in Psychology*, 8, 558. <https://doi.org/10.3389/fpsyg.2017.00558>
- Valian, V. (2015). Bilingualism and cognition. *Bilingualism*, 18(1), 3–24. <https://doi.org/10.1017/S1366728914000522>
- Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition*, 137, 40–46. <https://doi.org/10.1016/j.cognition.2014.12.008>
- von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, 145(2), 246. <https://doi.org/10.1037/xge0000120>
- Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, 20(3), 188–196. <https://doi.org/10.1016/j.lindif.2010.02.004>
- Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, 78(3), 332–348. <https://doi.org/10.1006/brln.2001.2474>
- Wang, M., Perfetti, C. A., & Liu, Y. (2005). Chinese–English biliteracy acquisition: Cross-language and writing system transfer. *Cognition*, 97(1), 67–88. <https://doi.org/10.1016/j.cognition.2004.10.001>
- Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese Listeners to Perceive Thai Tones: A Preliminary Report. *Language Learning*, 54(4), 681–712. <https://doi.org/10.1111/j.1467-9922.2004.00283.x>
- Werker, J. F. (1986). The effect of multilingualism on phonetic perceptual flexibility. *Applied Psycholinguistics*, 7(2), 141–155. <https://doi.org/10.1017/S0142716400007360>
- Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 25(1), 121–133. [https://doi.org/10.1016/S0163-6383\(02\)00093-0](https://doi.org/10.1016/S0163-6383(02)00093-0)
- Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, 106(7), 2468–2471. <https://doi.org/10.1073/pnas.0809035106>
- Wong, P. C. (2002). Hemispheric specialization of linguistic pitch patterns. *Brain research bulletin*, 59(2), 83–95. [https://doi.org/10.1016/S0361-9230\(02\)00860-2](https://doi.org/10.1016/S0361-9230(02)00860-2)
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420. <https://doi.org/10.1038/nn1872>
- 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(2), 123–139. <https://doi.org/10.1016/j.jml.2012.09.004>
- Yip, M. (2002). *Tone*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139164559>
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, 6(1), 37–46. [https://doi.org/10.1016/S1364-6613\(00\)01816-7](https://doi.org/10.1016/S1364-6613(00)01816-7)

Appendix I. Bilingual language background (apart from Dutch)

In this table, the other language spoken by the Dutch simultaneous bilingual participants were listed.

Language	Number of participants
Arabic	2
Danish	1
English	11
French	2
Frisian	3
German	3
Spanish	1
Tamil	1

Appendix II. Fo value in four points of the naturally-produced stimuli

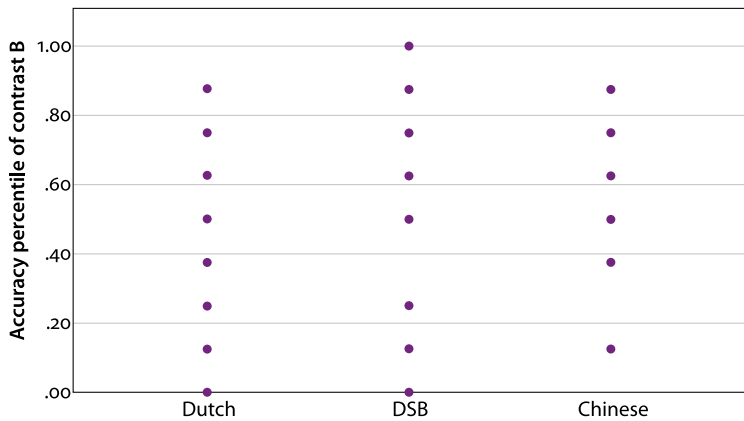
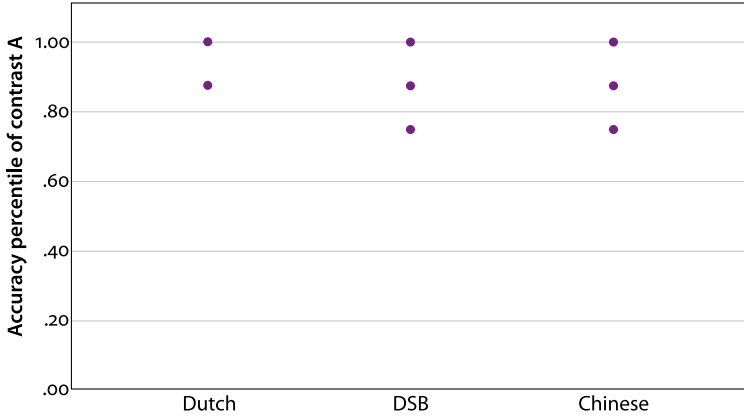
In this table, the four numbers (1–4) represent the four naturally produced contrasts (A) from which four manipulated contrasts (B) were created in PRAAT, resulting in eight pairs in total.

Pair	Contrast	Starting Point	Interpolate 1	Interpolate 2	Ending Point
1	A	288.30	289.00	289.00	290.00
	B	302.93	294.71	262.43	252.97
	B	324.87	303.29	222.57	197.43
	A	339.50	309.00	196.00	160.40
2	A	275.50	277.63	279.77	281.90
	B	291.66	287.25	255.64	247.28
	B	315.91	301.68	221.93	195.35
	A	332.08	311.30	198.80	160.73
3	A	278.80	276.67	274.53	272.39
	B	307.29	292.91	243.83	238.14
	B	350.04	317.26	197.78	186.75
	A	378.54	333.49	167.08	152.50
4	A	272.50	272.83	273.15	273.47
	B	303.05	284.98	245.03	240.13
	B	348.88	303.21	202.85	190.13
	A	379.43	315.37	174.73	156.79

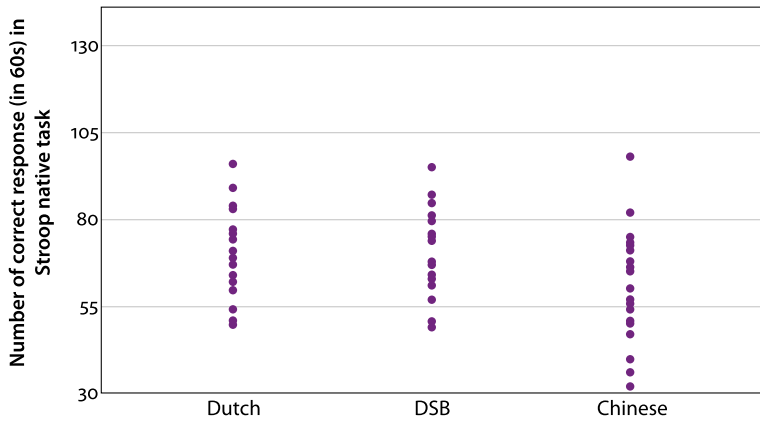
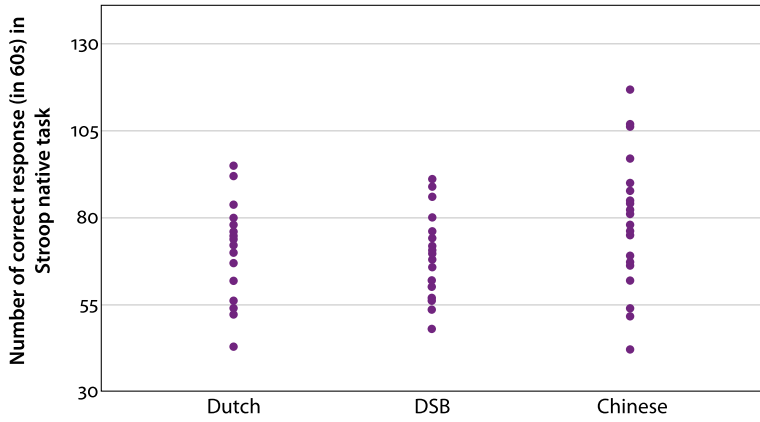
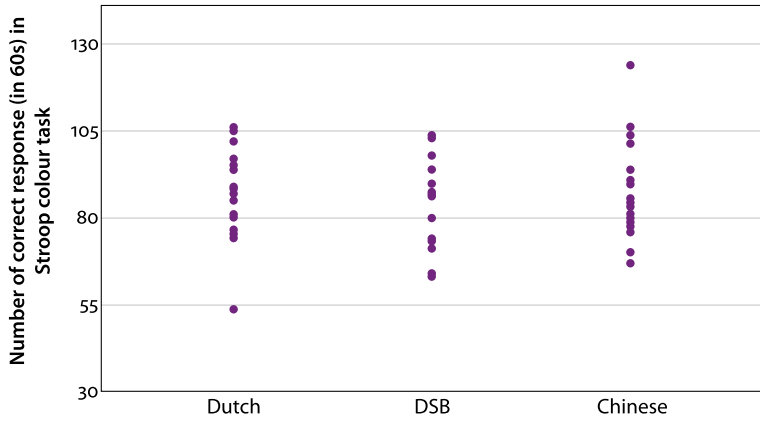
Appendix III. Scatter-plot across Experiments

Note: There are 24 participants per group. The number of dots may be less than 24 due to the overlap if participants share same scores.

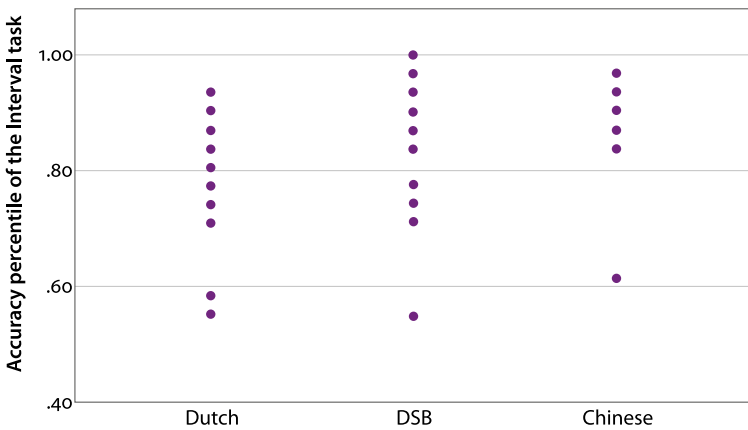
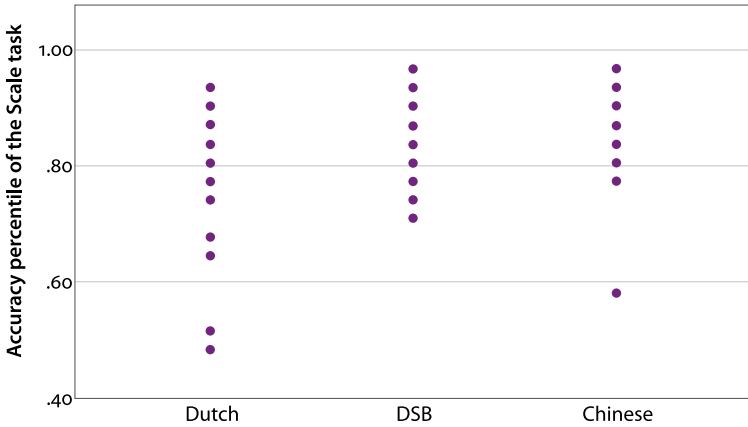
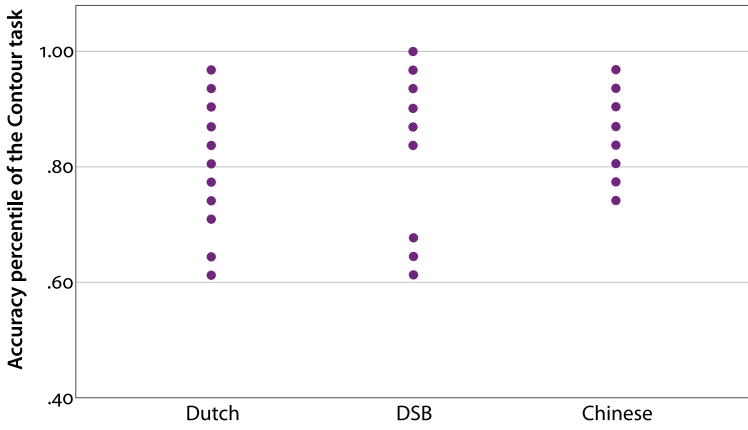
Experiment 1 Language perception, including two contrasts

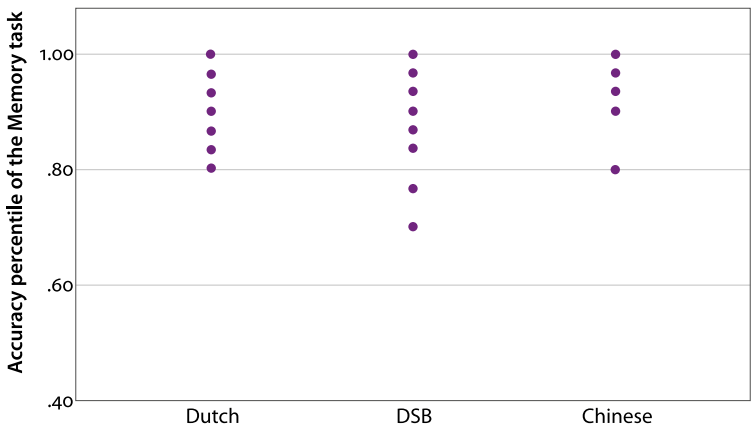
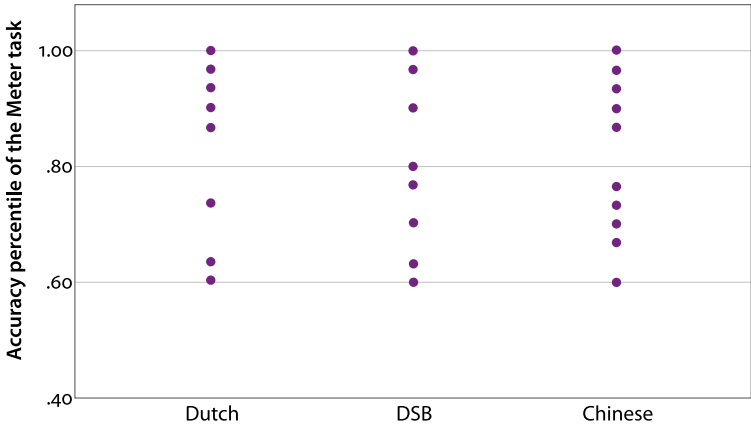
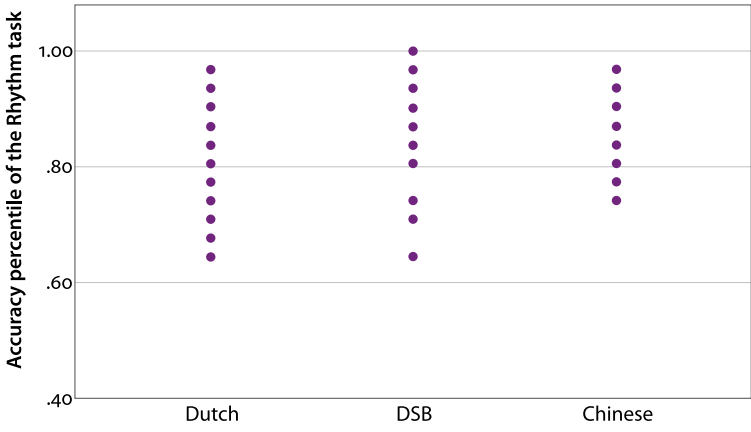


Experiment 2 Executive function, including three Stroop tasks, two Trail Making tasks, and one Digital Symbol Substitution task



Experiment 3 Music perception, including six tasks: contour, scale, interval, rhythm, meter, memory.





Address for correspondence

Liqun Liu
56 Second Avenue
Sydney 2747 Kingswood
Australia
l.liu@westernsydney.edu.au
 <https://orcid.org/0000-0001-8671-5098>

Co-author information

Ao Chen
Beijing Language and Culture University
Utrecht University
a.o.chen@uu.nl

René Kager
Utrecht University
r.w.j.kager@uu.nl
 <https://orcid.org/0000-0002-5811-839X>

Publication history

Date received: 14 May 2019
Date accepted: 17 September 2020
Published online: 19 November 2020