LANGUAGE EXPERIENCE INFLUENCES NON-LINGUISTIC PITCH PERCEPTION

Gang Peng
The Chinese University of Hong Kong
SIAT, Chinese Academy of Sciences

Diana Deutsch and Trevor Henthorn
University of California at San Diego

Danjie Su
University of California at Los Angeles

William S-Y. Wang
The Chinese University of Hong Kong
Academic Sinica, Taipei
University of California at Berkeley

ABSTRACT

Following publication of the Sapir-Whorf hypothesis, evidence has accumulated for the influence of language experience on perception. There are thousands of languages in the world which make use of pitch patterns to construct words much as vowels and consonants are used, among which Mandarin (a.k.a. Putonghua) is a typical tone language. This study examines the effect of language experience (tone language experience vs. nontone language experience) on non-linguistic pitch perception. First, we show a significantly higher prevalence of absolute pitch among native tone-language-speaking music students than among nontone-language-speaking music students. Moreover, we show that language experience shapes the perception of tone sweeps, extending the influence of language prototypes from the linguistic domain to the non-linguistic domain. Taken together, these results demonstrate that language experience affects auditory perception, and so provide evidence for the Sapir-Whorf hypothesis in the auditory modality.

SUBJECT KEYWORDS

Absolute pitch **T**one sweep **S**apir-Whorf hypothesis **P**itch perception **T**one language

1. INTRODUCTION

We begin our lives with the biological capacity to learn any language. By the time we are five or six, we are fully assimilated into English, Chinese, Swahili, or any of the several thousand languages that are currently spoken around the world. Each of these languages is a highly complex system, which represents the world in a unique way, with its own stock of sounds, words and phrases, and its own grammatical constructions. How do these different systems of representation influence our perception and behavior? This question was highlighted by two influential American linguists around the middle of the last century, and is referred to as the Sapir-Whorf hypothesis (Carroll 1956; Kay and Kempton 1984). However, because of the difficulty in stating the hypothesis in empirical testable terms, it has not so far led to much experimental research.

Recently, however, this hypothesis has sparked renewed attention, including the gathering of empirical evidence (Reines and Prinz 2009), especially in the realm of color perception (Drivonikou et al. 2007; Franklin et al. 2008; Gilbert et al. 2006; Kay et al. 2008; Roberson et al. 2007; Winawer et al. 2007). In a series of wide ranging experiments with adults, infants, aphasics, as well as two patients with callosotomy, it has been shown that color discrimination can be achieved faster by the left hemisphere, even when the subject responds by silently pressing keys (Kay et al. 2008). Furthermore, neuroimaging experiments have shown that the brain regions in the left hemisphere involved in such perceptual decisions are those that are also used in word finding processes (Siok et al. 2009; Tan et al. 2008). These experiments lead to the conclusion that our behavior in perceiving color is influenced by how the words of our language classify that portion of the electromagnetic continuum, in support of the Sapir-Whorf hypothesis. Moreover, support for lateralization of the Whorf effect has been found beyond the realm of color discrimination, where the perception of animal figures (cats and dogs) was more strongly affected by linguistic categories for stimuli presented to the right visual field than those presented to the left (Gilbert et al. 2008). Hence the term "Lateralized Whorf" is sometimes used to refer to these findings.

We now turn our attention from the visual modality to spoken language. Every language makes use of pitch to signify intonation patterns. However, a considerable portion of languages additionally use pitch to contrast lexical meanings, i.e., tone languages (Wang 1967, 1972, 1973). The various varieties of Chinese are all tone languages, among which Mandarin and Cantonese are two well-known examples (Peng 2006). In this regard, Cheng (1973) gives an overview of tone systems of Chinese varieties. Previous studies have demonstrated that the additional use of pitch for lexical contrasts shapes linguistic pitch perception (Hallé et al. 2004; Gandour 1983; Lee et al. 1996; Peng et al. 2010; Wang 1976; Xu et al. 2006). Recently, a series of studies have shown that language experience also shapes non-linguistic pitch perception; i.e., the prevalence of absolute pitch (AP: The ability to label or produce musical tones in the absence of a reference tone) among musicians, with native tone-language-speaking musicians having a significantly greater prevalence of AP (Deutsch 2002; Deutsch et al. 1999, 2004a, 2006, 2009).

In the present study, we continued to explore this issue in the auditory modality with two extensions: First, we investigated the prevalence of AP in a very larger number of students in music departments at regular universities in South China (these were native Mandarin speakers, and constituted the tone language group), and compared the AP performance of these music students with that of music students at a regular university in the U.S (these were English speakers, and constituted the nontone language group). We also compared these results with earlier published findings for nontone language speakers at a music conservatory in the U.S. , for which the sample size was also very large (Deutsch et al. 2009). The present study extended this investigation to students who were not in music conservatories, but rather in regular universities, so that our findings are more representative of the general population of musically literate individuals. As another issue, it has been suggested that fixed-do systems of musical training are more conducive to the development of AP than are moveable-do systems. In fixed-do systems, solfege symbols (do, re, mi, etc.) are used to define actual pitches. In contrast, in moveable-do systems the terms (C, C#, D, etc.) are used to define actual pitches, and the terms (do, re, mi etc.) are used instead to define the roles of pitches relative to a tonic. It has been surmised that since AP is more prevalent in some countries where fixed-do training is common, such as Japan, and is rare in other countries, such as the U.S., where moveable-do training is more common, then fixed-do training enhances the predisposition to acquire AP (Gregersen et al. 2001). However, AP is also rare in other countries, such as France, where fixed-do training is again common, so the argument in favor of fixed-do training based on prevalence of AP in a few selected countries is a problematic one. To investigate the effect in our Chinese population, the subjects had the option of responding using either fixed-do or moveable-do terminology, so indicating the type of training they had received.

Second, we examined the influence of language experience on categorizing nonspeech tone sweeps which had the same fundamental frequency (F0) distribution as their speech counterparts (Peng et al. 2010), where the two endpoints of the rising continuum represented the Mandarin high level tone and high rising tone respectively. By using the above two experiments, we investigated how language experience influences the perception of non-linguistic familiar sounds, such as musical tones by musicians, and non-linguistic unfamiliar sounds, such as tone sweeps by nonmusicians.

2. METHODS

2.1 Musical tone labeling task

2.1.1 Participants

The tone language (native Mandarin) group consisted of 298 subjects (70 males and 228 females) who were first- or second-year students at music departments of 3 universities in Guangzhou, P.R. China: With 144 subjects from South China Normal University, 60 subjects from Guangdong University of Foreign Studies, and 94 subjects from South China University of Technology, with a mean age of 20.2 (range 17-27) years. Among the above 298 subjects, there were 15 subjects who reported having had no musical training. Consequently, only data from the remaining 283 subjects were analyzed. The nontone language (native English) group consisted of 24 subjects (14 males and 10 females), who were first- or second-year students at University of California, San Diego (UCSD) with a mean age of 26.3 (range 22-30, with one outlier of 50) years. In both groups, all students who were invited to take the test of AP agreed to do so; there was therefore no self-selection of subjects within either group.

2.1.2 Stimuli

The stimuli were a set of 36 notes that spanned the three-octave range from C3 (131 Hz) to B5 (988 Hz). These notes were piano tones generated on a Kurzweil K2000 synthesizer which was tuned to the standard A4 of 440 Hz, and were 500 ms in duration.

2.1.3 Procedure

Subjects were administered a test for AP that was identical to that given in (Deutsch et al. 2006, 2009). Tones were presented in three blocks of 12, with 4.25 s intervals between onsets of notes within a block, and 39-s rest periods between blocks. The subjects were asked to write down the name of each tone (C, C#, D, and so on, or “do”, “do sharp”, “re” in Chinese notation style) when they heard it. In order to minimize the use of relative pitch as a cue, all successively presented tones were separated by an interval larger than an octave.

The three test blocks were preceded by a practice block in which four successive tones were presented. The subjects were given no feedback, either during the test blocks or during the practice block. The notes were played to subjects via a CD or DVD player, amplifier, and two loudspeakers. The subjects were also asked to fill out a questionnaire concerning their music education, where they and their parents had lived, and the languages they and their parents spoke.

2.2 Tone sweep labeling task[1]

2.2.1 Participants

Eighteen native speakers of Mandarin (tone) at Peking University in Beijing (8 females and 10 males, mean age = 23.2 yrs, SD = 2.3), and 15 native speakers of German (nontone) at the Max Plank Institute in Leipzig in Germany (7 females and 8 males, mean age = 28.3, SD = 4.8) were recruited for the tone sweep labeling task. No participant reported formal musical training, or any speech, language, or hearing difficulty.

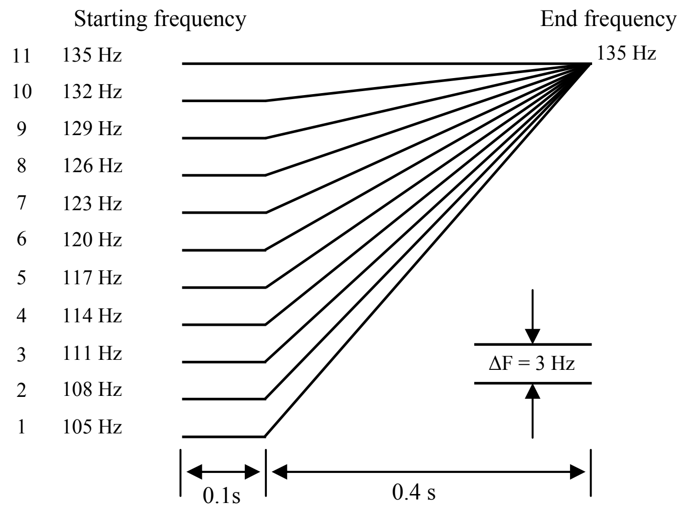

Fig. 1. Schematic diagram of the pitch contours of the 11 tone sweeps (adapted from Wang 1976).

2.2.2 Stimuli

Fig. 1 shows a schematic diagram of the tone sweeps used in this task. The 11 tone sweeps were created by software Praat (Version 5.1.05. Retrieved on May 1, 2009, from http://www.praat.org/). The detailed procedures for constructing the tone sweeps have been described in Peng et al. (2010). Complex tones have been used in the literature (e.g. Xu et al. 2006) as nonspeech stimuli, but sinusoidal tone sweeps (without harmonic structure) were used in the current study because they differ more than do complex tones from their speech counterparts.

2.2.3 Procedure

Two sounds (the two endpoints of the continuum) were played to participants multiple times (3–8 times according to participants' request) at the beginning of the test, and the participants were asked to remember these two representative sounds (the stimulus Number 11 was defined as 'Sound 1', while the stimulus Number 1 was defined as 'Sound 2'). The stimuli in the continuum were presented to the participants in random order, and the participants were asked to press key '1' when they heard

'Sound' 1 and to press key '2' when they heard 'Sound 2' (two-alternative forced choice, 2AFC), categorizing each stimulus into either category represented by 'Sound 1' or 'Sound 2'. Once a response was collected or 3500 ms had elapsed from the onset of the stimulus, whichever came first, the next stimulus was presented automatically following a 300 ms pause. The 11 stimuli were repeated twice in a block, with 5 such blocks in total. There was an additional practice block before the five test blocks.

3. RESULTS

3.1 Labeling Musical tones by musicians

3.1.1 Effect of language background

Before data analysis, three age groups were defined according to a previous study (Deutsch et al. 2009) in order to make better comparison: ages 2--5; ages 6--9; and ages ≥10. Fig. 2. shows the percentage correct responses on the test of AP, as a function of age of onset of musical training. A much higher score was produced by the tone language group for age-of-onset groups 2--5, and 6--9, with clear performance decrements for age-of-onset group ≥ 10.

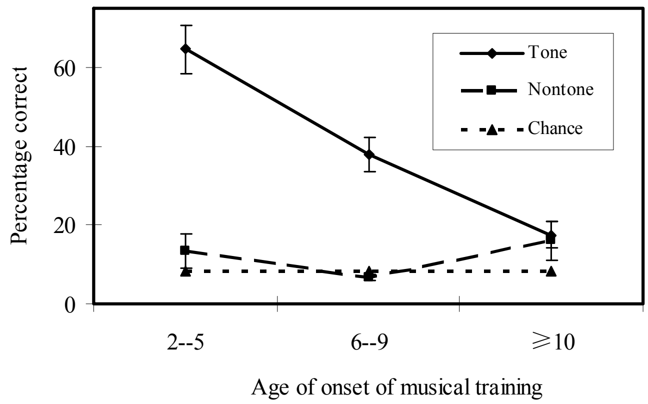

Fig. 2. Percentage correct responses on the test of AP, as a function of age of onset of musical training. The error bars represent standard error. The solid line shows the performance of the tone language group. The dashed line shows, as a comparison, the performance of native speakers of nontone language group. The short dashed line labeled chance represents chance performance on the task.

Statistical tests were carried out to determine the significance of the observation. A 2 × 2 × 3 analysis of variance (ANOVA) test was performed, with *language group* (tone vs. nontone), *gender* (male vs. female), and *age-of-onset group* (age of onset of musical training: 2--5 vs. 6--9 vs. ≥ 10) as factors. A highly significant effect of *language group* was found [$F(1, 295) = 15.47$, $p < .001$]. A highly significant effect of *age group* was also found [$F(2, 295) = 6.82$, $p < .001$]. The effect of *gender* was found to be non-significant [$F(1, 295) = .21$, $p = .64$]. The interaction between *language group* and *age group* was highly significant [$F(2, 295) = 7.55$, $p < .001$].

Further analyses examined the effect of age of onset of training for each language group separately. For nontone language group, this effect was non-significant [$F(2, 21) = 1.76$, $p = .20$], as would be expected from the small sample size combined with the overall very low performance. Previous findings (cf. Deutsch et al. 2006, 2009) have all pointed to an effect of age of onset of musical training for nontone language speakers. However, for the tone language group, the effect of age of onset of training was found to be highly significant [$F(2, 280) = 56.68$, $p < .001$]. Post-Hoc Bonferroni comparisons revealed that the effects of age of onset were found to be significant for each pair of age-of-onset groups of the tone language group ($p < .001$). The effect of *language group* was found to be significant for the age group 2--5 [$F(1, 42) = 8.12$, $p < .01$], significant for the age group 6--9 [$F(1, 94) = 6.63$, $p < .05$], but non-significant for the age group ≥ 10 [$F(1, 165) = .10$, $p = .75$].

3.1.2 Effect of type of response

Fig. 3. shows the percentage correct responses on the test of AP, as a function of age of onset of musical training and type of response. For each age group, the performance of the subjects among those who had received moveable-do training was higher than that of subjects among those who had received fixed-do training.

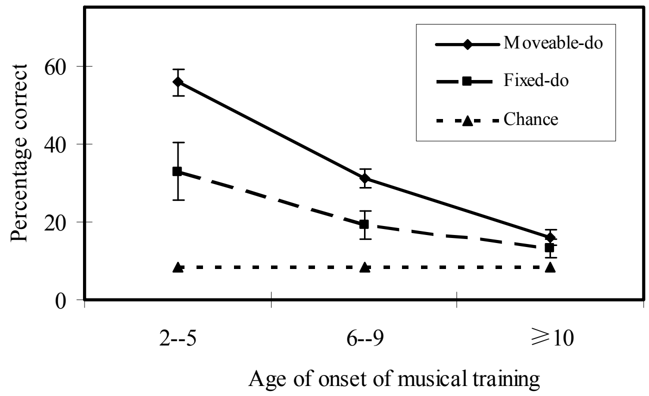

Fig. 3. Percentage correct responses on the test of AP, as a function of age of onset of musical training. The error bars represent standard error. The solid line shows the performance of the group responding with moveable-do terminology. The dashed line shows the performance of the group responding with fixed-do terminology. The short dashed line labeled chance represents chance performance on the task.

Statistical tests were carried out to determine the significance of the observation. A 2 × 3 ANOVA test was performed, with *type of response* (fixed-do vs. moveable-do) and *age-of-onset group* as factors. A highly significant effect of *type of response* was found [$F(1, 277) = 15.05$, $p < .001$]. A highly significant effect of *age group* was again found [$F(2, 277) = 27.33$, $p < .001$]. The interaction between *type of response* and *age group* was significant [$F(2, 277) = 4.46$, $p < .05$].

Further analyses examined the effect of *type of response* for each age group separately. The effect of *type of response* was found to be significant for the age group 6--9 [$F(1, 85) = 4.55$, $p < .05$], but non-significant for the age group 2--5 [$F(1, 36) = 3.02$, $p = .09$] and age group≥ 10 [$F(1, 156) = 2.66$, $p = .11$]. For the age group 2--5, though the result did not reach significance, the performance for the moveable-do group was obviously better than that of the fixed-do group. This non-significant result was likely due to the small number of subjects (only 7 subjects) for the fixed-do group.

3.2 Labeling Tone sweeps by nonmusicians

Concerning the experiment on tone sweeps, Fig. 4 depicts the response curves for a Mandarin subject. The subject labeled stimuli Number 1, 2, and 3 with 100% as "Sound 2", and stimuli Number 7, 8, 9, 10, and 11 with 0% as "Sound 2", i.e., with 100% as "Sound 1". Boundary position and boundary width of response curves were assessed by Probit analyses of individual response curves (Finney 1971): The boundary position was defined as the 50% crossover point, and the boundary width was defined as the linear distance between the 25th and 75th percentiles of either fitted response curve as determined by the mean and standard deviation obtained from Probit analysis.

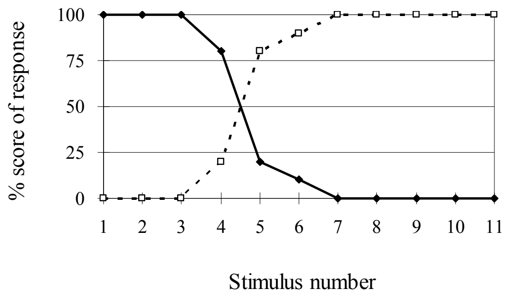

Fig. 4. The solid curve represents the response curve for "Sound 2", and the dotted line represents the response curve for "Sound 1", for a Mandarin subject. The sum of percentiles of the responses to one stimulus is 100, since it is a two-alternative forced choice task.

In several recent studies, the boundary position was found to be language independent, while the boundary sharpness or width was found to be language dependent (Xu et al. 2006; Peng et al. 2010). Therefore, the main focus was on boundary width here. Fig. 5 shows the distributions of the boundary widths for the two language groups. There were two extreme outliers for the tone group, and one for the nontone group, which fell outside the upper bound of the *y* axis of Fig. 5. Moreover, these outliers were also out of the range of corresponding *mean* ± 2 × *standard deviation*. Therefore these three outliers were discarded in further analyses.

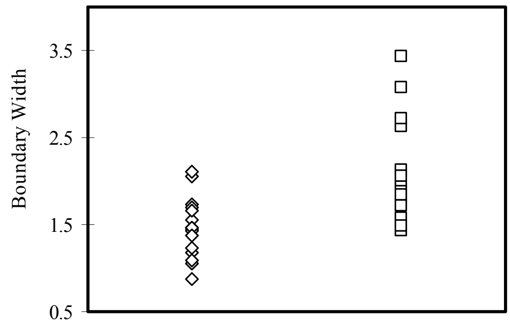

Fig. 5. Distribution of boundary width on the test of tone sweep categorization, as a function of language. Each point represents the boundary width for one subject. Diamonds represent widths for tone language speakers, and squares for nontone language speakers.

As expected, non-significant effect of *language* (tone vs. nontone) on boundary position was found [$F(1, 28) = .04$, $p = .85$]. A highly significant effect of *language* on boundary width was found [$F(1, 28) = 13.71$, $p < .001$], with marginal mean of 1.46 for the tone group, and 2.12 for the nontone group. The effect of *gender* on boundary width was found to be non-significant [$F(1, 28) = 1.734$, $p = .20$].

4. GENERAL DISCUSSION

The findings regarding the influence of language experience, and age of onset of musical training, on the prevalence of AP are highly consistent with findings in the previous literature (Deutsch et al. 2006, 2009). The performance level on the AP test was significantly higher among those who were native tone language speakers compared with those who were nontone language speakers, indicating that the use of pitch as lexical contrasts facilitates the development of AP. Although the number of nontone language subjects in the present study was small, the data present a striking replication of the University of Southern California (USC) results reported earlier (Deutsch et al. 2009), where data from 176 nontone language subjects were analyzed. Given the current very large scale of the AP tests for tone language subjects (283 effective subjects), together with the striking replication of the results for the nontone group, the current findings are presumably representative for the target populations. The effect of age of onset of training was again consistent with previous research (Deutsch et al. 2006, 2009): The earlier the onset of musical training the higher the probability of acquiring AP ability. More specifically, AP was found to be most prevalent among those who had begun musical training at ages 2--5, less prevalent at ages 6--9, and very rare at or after age 10.

As shown in Fig. 2, for the tone language group, the first significant performance decrement from 51.5% to 27.5% indicates that the best period for acquiring AP is ages 2--5, and the second significant decrement from 27.5% to 14.9% indicates that successfully acquiring AP is quite unusual at age 10 or later. However, for the nontone language group, together with reference to the data reported in (Deutsch et al. 2006, 2009) where a much larger number of nontone subjects were tested, AP performance was about chance level when the age of onset of musical training was greater than 6, which is consistent with the sensitive age limit (5-6 years) of the early learning hypothesis (Vitouch 2003). This finding resonates with the claim that the prevalence of AP is extremely rare in the U.S. and Europe, with an estimate of less than one in 10,000 in the general population (Profita and Bidder 1988). It also strengthens the claim that for tone language speakers the acquisition of absolute pitch is akin to acquiring a second tone language (Deutsch et al. 2004a). For tone language speakers, the neural circuitry for acquiring the association between pitches and verbal labels is already in place, so that they are able to build the ability of acquiring absolute pitch on it with greater success. Moreover, better pitch processing ability as a result of music training also carries over to speech processing (Wong et al. 2007; Oechslin et al. 2010).

With respect to the second experiment, the boundary width of the tone sweep labeling task is significantly narrower for the tone language group, i.e., it shows a sharper transition in this group than in the nontone language group. The high level tone, with pitch contour represented as stimulus Number 11, and the high rising tone, with pitch contour represented as stimulus Number 1, both exist in the phonological system of Mandarin.

Our findings with respect to tone sweeps are in accordance with existing theoretical frameworks and associated research. The perceptual magnet effect hypothesizes that native language prototypes pull neighboring speech sounds toward them (Kuhl et al. 2008), which would make the labeling transition sharper for tone language speakers. This hypothesis is tightly connected with evidence of categorical perception for native speech sounds (Liberman 1996; Liberman and Whalen 2000). Further, the transition boundary may also be associated with the idea of footprint in the quantal theory of speech perception (Stevens 1998, 2000). Our findings indicate that the influence of the two language prototypes, i.e., the high level tone and the high rising tone, extends to perception in the nonspeech domain. Besides its influence at the behavioral level, language experience has been found to shape the perception of tonal contours in nonspeech contexts at the level of the brainstem (Krishnan et al. 2008).

As described earlier, we found that native speakers of tone languages performed significantly better than a matched group of nontone language speakers in our test for absolute pitch. Correspondingly, Fig. 5 shows that native speakers of tone languages without formal musical training show a significantly narrower categorization boundary, when compared with a nearly matched pool of native speakers of nontone languages. The stimuli for the latter experiment would have been familiar to the musician subjects, but not to the nonmusicians. Yet in both cases it is clear that whether the native language is tonal or not has a highly significant effect on their pitch perception, which provides evidence for the Sapir-Whorf hypothesis in the auditory modality. Further, the performance of the subjects among those who had received moveable-do training was higher than that of subjects among those who had received fixed-do training, especially for the subjects who started musical training before age 10. These findings show that fixed-do training cannot account for the high level of performance among the Chinese subjects in our study.

The basis for AP is still unclear. An enhanced leftward asymmetry of the planum temporale has been found among Western AP possessors (Keenan et al. 2001; Schlaug et al. 1995; Zatorre 2003; Zatorre et al. 1998), and this might indicate a hardwired, and so possibly genetic basis of AP among nontone language speakers. However, in the study by Deutsch et al. (2009) it was found that, among subjects of the same ethnic heritage, significant differences in performance level emerged depending on degree of fluency in speaking a tone language; moreover, the performance level of subjects of East Asian ethnicity who were not fluent in speaking a tone language did not differ significantly from that of subjects of Caucasian ethnicity (who also did not speak a tone language). These two findings indicate that there is no genetic advantage to East Asian ethnicity for the predisposition to acquire AP. Nevertheless, the factors determining the acquisition of AP by nontone language speakers still remain unresolved.

A recent paper raised the question "Is there an Asian advantage for pitch memory?", and answered it in the negative (Schellenberg and Trehub 2008). However there were important differences between their subject characteristics and ours. In their study, the subjects (mean age 10 years, 9 months) were not potential musicians, since the average amount of musical training they had obtained was only 17 months. In contrast, in the present study, the average amount of musical training obtained before age 11 for those who began musical training no later than 9 years old was more than 5 years. Therefore, our findings do not conflict with those of Schellenberg and Trehub, while at the same time they do show an Asian advantage.

Pitch ability in language and music have been shown to be closely related (Deutsch et al. 2004a; Deutsch et al. 2004b; Jackendoff and Lerdahl 2006; McMullen and Saffran 2004), even since early childhood (Chen-Hafteck 1998; Garfias 1990). In addition to possibly shared neural circuitry at the cortical level, previous research has indicated a common subcortical manifestation also (Wong et al. 2007). No doubt early development in a tone language environment sharpens the acuity of pitch perception in language, and this advantage is likely to be carried over to the nonspeech domain, e.g. musical notes, or tone sweeps.

We can hypothesize that the advantage for AP possession among tone language speakers may well begin during fetal life. Native speakers of tone languages are exposed to tonal contrasts even before birth (Holst et al. 2005; Welling 2010), and 2-3 day-old infants are able to process pitch intervals (Stefanics et al. 2009). Native speakers of nontone languages do not have pitch training comparable to that of native speakers of tone languages before they begin music lessons. Nevertheless, it has been found that 8-month infants from nontone language speaking families were more likely to track patterns of AP than of relative pitches, while adults relied primarily on relative pitch (Saffran and Griepentrog 2001). Since some speech abilities appear to be lost roughly 9 months after birth if the they are not reinforced by environment (Kuhl 2004), the pitch processing ability of nontone language speakers might well be reduced if they do not receive alternative pitch training, such as in music.

So far, the most widely used contrast for language experience in this regard has been a general one; i.e., tone language vs. nontone language, and only a limited number of studies have considered contrasts in different tone systems. Peng et al. (2010) found that Cantonese and Mandarin subjects showed different discrimination patterns for speech tone continua, indicating that experience with different tone languages differentially modulate pitch perception. It is well known that Cantonese has a rich tone inventory in comparison to Mandarin (Peng 2006). It is natural to ask whether the richer tone inventory of Cantonese provides its users with better pitch ability in general. Moreover, Cantonese has multiple level tones, so processing pitch height must be particularly useful in distinguishing these tones. Such a heavy dependence on pitch height in Cantonese may enhance the ability of its speakers to identify piano music notes, which are also level in pitch. Based on this consideration, we further hypothesize that the prevalence of absolute pitch among native Cantonese musicians might be even greater than that among native Mandarin musicians. This is a fascinating topic that merits further study.

ACKNOWLEDGEMENTS

The work described in this paper was partially supported by grants from the Shun Hing Institute of Advanced Engineering of The Chinese University of Hong Kong (Project No. BME-8115020), from National Natural Science Foundation of China (NSFC: 11074267), and from the Research Grant Council of Hong Kong (Project No. 455911). We thank South China Normal University, Guangdong University of Foreign Studies, and South China University of Technology for their support.

NOTES

1. The original data that were obtained on this task were published in Peng et al. (2010). Since the aim of the present study was to examine how nonmusicians label tone sweeps, in relation to their linguistic background, participants with professional music training were discarded from the present analysis, and only the data from the remaining 18 Mandarin and 15 German participants were here analyzed.

REFERENCES

CARROLL, John B. 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge, Mass.: Massachusetts Institute of Technology Press.

CHENG, Chin-Chuan. 1973. A quantitative study of Chinese tones. *Journal of Chinese Linguistics* 1:93–110.

CHEN-HAFTECK, Lily. 1998. Pitch abilities in music and language of Cantonese-speaking children. *International Journal of Music Education* 31:14–24.

DEUTSCH, Diana. 2002. The puzzle of absolute pitch. *Current Directions in Psychological Science* 11:200–204.

DEUTSCH, Diana, K. Dooley, T. Henthorn and B. Head. 2009. Absolute pitch among students in an American music conservatory: Association with tone language fluency. *Journal of the Acoustical Society of America* 125:2398–2403.

DEUTSCH, Diana, T. Henthorn and M. Dolson. 1999. Absolute pitch is demonstrated in speakers of tone languages. *Journal of the Acoustical Society of America* 106:2267.

DEUTSCH, Diana, T. Henthorn and M. Dolson. 2004a. Absolute pitch, speech, and tone language: Some experiments and a proposed framework. *Music Perception* 21:339–356.

_____. 2004b. Speech patterns heard early in life influence later perception of the tritone paradox. Music Perception 21:357–372.

DEUTSCH, Diana, T. Henthorn, E. Marvin and H.-S. Xu. 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.

DRIVONIKOU, Gilda Vasiliki, P. Kay, T. Regier, R.B. Ivry, A.L. Gilbert, A. Franklin and I.R.L. Davies. 2007. Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences of the United States of America* 104:1097–1102.

FINNEY, David John. 1971. *Probit Analysis* (3rd ed). Cambridge: Cambridge University Press.

FRANKLIN, Anna, G.V. Drivonikou, L. Bevis, I.R.L. Davies, P. Kay and T. Regier. 2008. Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences of the United States of America* 105:3221–3225.

GANDOUR, Jackson T. 1983. Tone perception in Far Eastern languages. *Journal of Phonetics* 11:149–175.

GARFIAS, Robert. 1990. Thoughts on the processes of language and music acquisition. In *Music and Child Development,* ed. Franz L. Roehmann and Frank R. Wilson, 100–105. St. Louis, MO: MMB Music, Inc.

GILBERT, Aubrey L., T. REGIER, P. KAY and R.B. IVRY. 2006. Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America* 103:489–494.

_____. 2008. Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain and Language* 105:91–98.

GREGERSEN, Peter K., E. Kowalsky, N. Kohn and E. W. Marvin. 2001. Early childhood music education and predisposition to absolute pitch: Teasing apart genes and environment. *American Journal of Medical Genetics* 98:280-282.

HALLÉ, Pierre A., Y.-C. Chang and C.T. Best. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics* 32:395–421.

HOLST, Manuela, H. Eswaran, C. Lowery, P. Murphy, J. Norton and H. Preissl. 2005. Development of auditory evoked fields in human fetuses and newborns: A longitudinal MEG study. *Clinical Neurophysiology* 116:1949–1955.

JACKENDOFF, Ray. and F. Lerdahl. 2006. The capacity for music: What is it, and what's special about it? *Cognition* 100:33–72.

KAY, Paul and W. Kempton. 1984. What is the Sapir-Whorf hypothesis? *American Anthropologist* 86:65–79.

KAY, Paul, T. Regier, A.L. Gilbert and R.B. Ivry. 2008. Lateralized Whorf: Language influences perceptual decision in the right visual field. In *Language, Evolution, and the Brain*, ed. James W. Minett and William S-Y. Wang, 261–284. Hong Kong: City University of Hong Kong Press.

KEENAN, Julian Paul, V. Thangaraj, A. Halpern and G. Schlaug. 2001. Planum temporale and absolute pitch. *NeuroImage* 14:1402–1408.

KRISHNAN, Ananthanarayan, J. Swaminathan and J.T. Gandour. 2008. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience* 21:1092–1105.

KUHL, Patricia K. 2004. Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5:831–843.

KUHL, Patricia K., B.T. Conboy, S. Coffey-corina, D. Padden, M. Rivera-gaxiola and T. Nelson. 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 363:979–1000.

LEE, Yuh-Shiow, D.A. Vakoch and L.H. Wurm. 1996. Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research* 25:527–542.

LIBERMAN, Alvin M. 1996. *Speech: A Special Code*. Cambridge, MA: MIT Press.

LIBERMAN, Alvin M. and D.H. Whalen. 2000. On the relation of speech to language. *Trends in Cognitive Sciences* 4:187–196.

MCMULLEN, Erim M.C. and J.R. Saffran. 2004. Music and language: A developmental comparison. *Music Perception* 21:289–311.

OECHSLIN, Mathias S., M. Meyer and L. Jäncke. 2010. Absolute pitch-functional evidence of speech-relevant auditory acuity. *Cerebral Cortex* 20:447–455.

PENG, Gang. 2006. Temporal and tonal aspects of Chinese syllables: A corpus-based comparative study of Mandarin and Cantonese. *Journal of Chinese Linguistics* 34(1):134–154.

PENG, Gang, H.-Y. Zheng, T. Gong, R.-X. Yang, J.-P. Kong and W. S-Y. Wang. 2010. The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics* 38:616–624.

PROFITA, Joseph and T.G. Bidder. 1988. Perfect pitch. *American Journal of Medical Genetics* 29:763–771.

REINES, Maria Francisca and J. Prinz. 2009. Reviving Whorf: The return of linguistic relativity. *Philosophy Compass* 4(6):1022–1032.

ROBERSON, Debi, H.J. Pak and R. Hanley. 2007. Categorical perception of colour in the left and right visual field is verbally mediated: evidence from Korean. *Cognition* 107:752–762.

SAFFRAN, Jenny R. and G.J. Griepentrog. 2001. Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology* 37:74–85.

SCHELLENBERG, E. Glenn and S. E. Trehub. 2008. Is there an Asian advantage for pitch memory? *Music Perception* 25:241–252.

SCHLAUG, Gottfried, L. Jaencke, Y. Huang and H. Steinmetz. 1995. In vivo evidence of structural brain asymmetry in musicians. *Science* 267:699–701.

SIOK, Wai-Ting, P. Kay, W.S.Y. Wang, A.H.D. Chan, L. Chen, K.-K. Luke and L.-H. tan. 2009. Language regions of brain are operative in color perception. *Proceedings of the National Academy of Sciences of the United States of America* 106:8140–8145.

STEFANICS, Gábor, G.P. Háden, I. Sziller, L. Balázs, A. Beke and I. Winkler 2009. Newborn infants process pitch intervals. *Clinical Neurophysiology* 120:304–308.

STEVENS, Kenneth N. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.

_____. 2010. Quantal theory, enhancement and overlap. *Journal of Phonetics* 38:10–19.

TAN, Li-Hai, A.H.D. Chan, P. Kay, P.-L. Khong, L.K.C. Yip and K.-K. Luke. 2008. Language affects patterns of brain activation associated with perceptual decision. *Proceedings of the National Academy of Sciences of the United States of America* 105:4004–4009.

VITOUCH, OLIVER. 2003. Absolutist models of absolute pitch are absolutely misleading. *Music Perception* 21:111–117.

WANG, William Shi-Yuan. 1967. Phonological features of tone. *International Journal of American Linguistics* 33:93–105.

_____. 1972. The many uses of F0. In *Linguistics and Phonetics to the Memory of Pierre Delattre*, ed. A. Valdman, 487–503. The Hague: Mouton.

_____. 1973. The Chinese language. *Scientific American* 228:53–62.

_____. 1976. Language change. *Annals of the New York Academy of Sciences* 208:61–72.

WELLING, Deborah R. 2010. Hearing and language development. In *Language Development: Foundations, Processes, and Clinical Applications,* ed. B.B. Shulman and N.C. Capone. 95–134. Sudbury, Mass: Jones and Bartlett Publishers.

WINAWER, Jonathan, N. Witthoft, M.C. Frank, L. Wu, A.R. Wade and L. Boroditsky. 2007. Russian blues reveal the effects of language on color discrimination. *Proceedings of the National Academy of Sciences of the United States of America* 104:7780–7785.

WONG, Patrick, E. Skoe, N.M. Russo, T. Dees and N. Kraus. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience* 10:420–422.

XU, Yisheng, J.T. Gandour and, A.L. Francis. 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America* 120:1063–1074.

ZATORRE, Robert J. 2003. Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience* 6:292–695.

ZATORRE, Robert J., D.W. Perry, C.A. Beckett, C.F. Westbury and A.C. Evans. 1998. Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proceedings of the National Academy of Sciences of the United States of America* 95:3172–3177.

语言经验影响非语言音高感知

彭刚
香港中文大学
中国科学院深圳先进技术研究院

苏丹洁
加州大学洛杉矶分校

王士元
香港中文大学
中央研究院, 台北
加州大学伯克莱分校

Diana Deutsch and Trevor Henthorn
加州大学圣地亚哥分校

题要
随着萨皮尔–沃尔夫假说的提出，语言经验影响感知已经累积了许多证据。世界上有几千种语言像利用元音和辅音一样利用音高模式来构词。其中，普通话就是一种典型的声调语言。本研究考察语言经验（声调语言经验对比非声调语言经验）对非语言音高感知的影响。首先，我们发现母语为声调语言的学习音乐的学生要比同等的母语为非声调语言的学习音乐的学生的绝对音高能力显著地好。此外，语言经验也影响对纯音扫描的感知，把语言经验对音高感知的影响力拓展到非语言且非熟悉范畴。总而言之，这些结果表明语言经验也影响一般听觉感知，进而为萨皮尔–沃尔夫假说提供了听觉模态方面的证据。

关键词
绝对音高 纯音扫描 萨皮尔–沃尔夫假说 音高感知 声调语言

Gang Peng 彭刚
Department of Linguistics and
Modern Languages
The Chinese University of Hong Kong
Shatin, New Territories, Hong Kong SAR
[gpeng@cuhk.edu.hk]