Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009) Jyväskylä, Finland
Jukka Louhivuori, Tuomas Eerola, Suvi Saarikallio, Tommi Himberg, Päivi-Sisko Eerola (Editors)

Absolute Pitch and Tone Language: Two New Studies

Diana Deutsch,[1] Jinghong Le,[2] Kevin Dooley[1] Trevor Henthorn,[1] Jing Shen[1], and Brian Head[3]

[1]*Department of Psychology., University of California, San Diego, U.S.A*
[2]*School of Psychology and Cognitive Science, East China Normal University, China*
[3]*Thornton School of Music, University of Southern California, U.S.A.*
[1]ddeutsch@ucsd.edu, [2]jhle@psy.ecnu.edu.cn, [3]kdooley@ucsd.edu,
[4]trevor@music.ucsd.edu, [5]jshen@psy.ucsd.edu, [6]head@usc.edu

ABSTRACT

Two new studies provide evidence in favor of the hypothesis that absolute pitch is strongly influenced by a speech-related critical period. The first study examined the prevalence of absolute pitch among students in at the University of Southern California's Thornton School of Music, as a function of age of onset of musical training, ethnicity, fluency in speaking a tone language, and country of early music education. Among those of East Asian ethnicity, performance on a test of absolute pitch was strongly correlated with fluency in speaking a tone language. The advantage of early onset of musical training did not interact statistically with the effects of tone language fluency, and further analyses showed that the results could not be explained by country of early music education. The second study investigated the pitch ranges of female speech in two relatively isolated villages in China. These pitch ranges clustered within each village, but differed significantly across the villages, indicating that, at least for speakers of tone language, the pitch range of speech is heavily influenced by an absolute pitch template that is developed through long term exposure to speech in the environment. Implications of these findings are discussed.

I. INTRODUCTION

Absolute pitch, defined as the ability to name or produce a note of given pitch in the absence of a reference note, is extremely rare in the U.S. and Europe, even among musicians (Bachem, 1955; Profita and Bidder, 1988). This rarity is perplexing, considering that musicians spend thousands of hours reading musical notes, playing the notes they read, and hearing the notes they play. The identification of pitches by absolute pitch nonpossessors occurs in a remarkably indirect way. To take the naming of colors as a comparison, when we identify a color as, say, 'red', we do not do this by viewing a different color whose name we have been given (say, 'blue') and then evaluating the relationship between the two colors. Yet nonpossessors of absolute pitch use just such an indirect process when identifying a note by name.

Deutsch (2002) suggested that, in nontone language speakers, the inability to label musical notes without the aid of a reference note stems from a lack of opportunity to form associations between pitches and their verbal labels during the critical period in which infants acquire other features of their native language. In contrast, infants who learn to speak a tone language (such as Mandarin, Cantonese, or Vietnamese) are provided with this opportunity: In tone languages, words assume entirely different meanings depending on the lexical tones in which they are enunciated, with tones being defined by both their pitch heights and their pitch contours. So, for example, in Beijing Mandarin the first tone is high and level, the second is mid-high and rising, the third is low and dipping, and the fourth is high and falling. When speakers of Mandarin hear the word 'ma' in the first tone and attribute the verbal label 'mother', and when they hear the word 'ma' in the third tone and attribute the verbal label 'horse', they using a pitch (or series of pitches) in the process of assigning verbal labels to sounds. And analogously, when absolute pitch possessors hear the note G# and attribute the label 'G#', or when they hear the note D and assign the label 'D' they are also associating pitches with verbal labels.

The hypothesis that absolute pitch is subject to a speech-related critical period received support from an experiment in which we asked native speakers of Vietnamese and Mandarin to recite the same list of words in their native language on two different days. We found that at least 1/3 of these subjects showed signed average pitch differences across days of less than ¼ semitone. A control group of speakers of English showed significantly less consistency across days in reciting a list of English words. We hypothesized that the greater pitch consistency of the tone language speakers resulted from their early acquisition of tone languages, so that they had learned to associate pitches with meaningful words in infancy (Deutsch et al., 2004).

Based on this line of reasoning, it was further conjectured that absolute pitch, which has traditionally been viewed as a musical faculty, originally evolved to subserve speech. It was further conjectured that when infants who acquire absolute pitch as a feature of their native language reach the age at which they can begin taking music lessons, they acquire absolute pitch for musical tones in the same way as they would acquire the tones of a second tone language. In contrast, children who had first acquired a nontone language would need to acquire the pitches of musical tones as though they were the tones of a first tone language.

This line of reasoning in turn led to the conjecture that the prevalence of absolute pitch for musical tones should be far higher among tone language speakers than among speakers of nontone languages such as English. To examine this conjecture, we documented the prevalence of absolute pitch in two large groups of music students. The first group were enrolled in a required course at the Central Conservatory of Music in Beijing; these all spoke Mandarin. The second group were enrolled in a required course at Eastman School of Music, and were all nontone language speakers. The two groups were divided into subgroups by age of onset of musical training. The subjects were given a test of absolute pitch that consisted of the 36 notes from the C below Middle C to the B almost three octaves above. On comparing the percentages of subjects who obtained a score of at least 85% on the test, we found that for all levels of age of onset of musical training, the percentage of those who met the criterion

was far higher among the tone language speakers than among the speakers of nontone language (Deutsch et al., 2006).

The above findings support the hypothesis that, if given the opportunity, infants can acquire absolute pitch as a feature of speech, which can then carry over into music. However, since the Mandarin speakers in the study of Deutsch et al. (2006) were all East Asian, and the nontone language speakers were all Caucasian, the findings could alternatively be interpreted as indicating a genetic basis for the capacity to acquire absolute pitch. As a further possible explanation, music education practices in China might be particularly conducive to the development of this ability. So in the first experiment to be described here, we examined the correlate with language in more detail, and also examined the alternative interpretation in terms of country of music education.

II. THE PREVALENCE OF ABSOLUTE PITCH AS A FUNCTION OF TONE LANGUAGE FLUENCY

In this study (Deutsch et al., 2009a), we explored the prevalence of absolute pitch among students at USC Thornton School of Music. There were 203 subjects in the study, 110 male and 93 female, with average age 19.5 years. 162 subjects were first- or second-year year students taking a required course at USC Thornton, and the remaining 41 subjects were members of the USC Community Orchestra. All those who were invited to take the test agreed to do so, so there was no self-selection of subjects for the experiment. The subjects filled out a detailed questionnaire concerning their music education, ethnic heritage, where they and their parents had lived, the languages they and their parents spoke, and how fluently they spoke each language (Deutsch et al., 2009a).

Based on their responses to the questionnaire we divided the subjects into several groups. Those in group *nontone* were Caucasian and stated that they spoke only nontone language fluently. The remaining subjects were all of East Asian descent, with both parents primarily speaking an East Asian tone language. We assigned these subjects to three further groups in accordance with their responses to the questionnaire. Those in group *tone very fluent* stated that they spoke an East Asian tone language 'very fluently'. Those in group *tone fairly fluent* stated that they spoke a tone language 'fairly fluently'. Those in group *tone nonfluent,* responded 'I can understand the language, but don't speak it fluently'.

We divided each group into two subgroups based on age of onset of musical training: One subgroup had begun musical training at ages 2-5, and the other at ages 6-9. There was also a third subgroup of nontone language speakers who had begun musical training at the age of 10 or later. However since all the subjects of East Asian descent had begun musical training by age 9, we confined our statistical analyses to the first two age groups.

We administered the same test of absolute pitch as we had given the subjects in our earlier study (Deutsch et al., 2006). It consisted of the 36 notes that spanned the three-octave range from the C below Middle C to the B almost three octaves above, and to minimize the use of relative pitch as a cue, all intervals between successive notes were larger than an octave. The notes were presented in three blocks of twelve, and the test was preceded by a practice block of four notes. No feedback was provided, either during the practice block, or during the test itself. The notes were piano samples that were generated on a Kurzweil synthesizer, and recorded on CD. The subjects listened to the CD through loudspeakers, and identified each note in writing.

Figure 1 shows, for each group and subgroup, the average percentage correct on the test for absolute pitch. The line labeled *chance* represents chance performance on the test. It can be seen that those subjects who stated that they spoke a tone language *very fluently* (shown in blue) showed remarkably high performance on the test. Indeed, their performance was far higher than that of the *nontone* language speakers (shown in black). It was also far higher than for the subjects of East Asian ethnicity but who did not speak a tone language fluently (i.e., the *tone nonfluent* group, shown in red). It was also higher than that of the *tone fairly fluent* speakers (shown in green), which was in turn higher than that of the *tone nonfluent* speakers, and of the *nontone* language speakers.

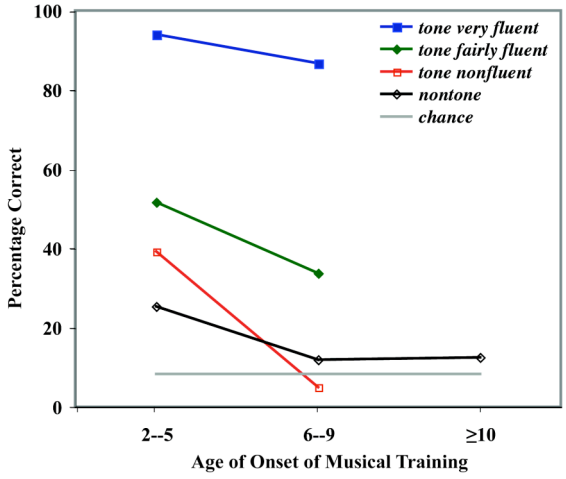

Figure 1. Percentage correct responses on the test of absolute pitch, for each group and subgroup. See text for details. Adapted from Deutsch et al., 2009a.

To make statistical comparison between the groups, we first carried out an overall test for an effect of gender, and found it to be nonsignificant. We then carried out a 4 x 2 ANOVA, with group and age of onset of musical training as factors. As expected from earlier studies (Baharloo et al., 1998; Deutsch et al., 2006) the effect of age of onset was significant, with those who had begun musical training at ages 2-5 performing at a higher level than those who had begun musical training at ages 6-9 [$F(1, 148)=5.15$, $p= 0.025$]. In addition, a highly significant effect of group was found [$F(3, 148)=35.43$, $p < 0.001$], which was overwhelmingly larger in absolute terms than the effect of age of onset of musical training. On post-hoc comparisons, the performance level of the *tone very fluent* group was significantly higher than that of the *nontone* group ($p < 0.001$), and was also significantly higher than that of the *tone nonfluent* group, as well as that of the *tone fairly fluent* group ($p < 0.001$). The performance level of the *tone fairly fluent* group was significantly higher than

that of the *nontone* group ($p = 0.003$), and also higher as a nonsignificant trend than that of the *tone nonfluent* group. The performance level of the *tone nonfluent* group did not differ from that of the Caucasian *nontone* group ($p > 0.05$). It is of particular importance to note that the performance of the (East Asian) *tone nonfluent* speakers was far poorer than that of the (East Asian) *tone very fluent* speakers, while at the same time not differing statistically from that of the (Caucasian) *nontone* language speakers.

We also performed a regression analysis taking only those subjects of East Asian ethnicity, and found that fluency in speaking a tone language was a highly significant predictor of performance ($R^2_{adj}=0.54$, $F(2,24)=16.09$, and $p < 0.001$). So these findings indicate that the differences in performance between the groups were determined by language rather than ethnicity.

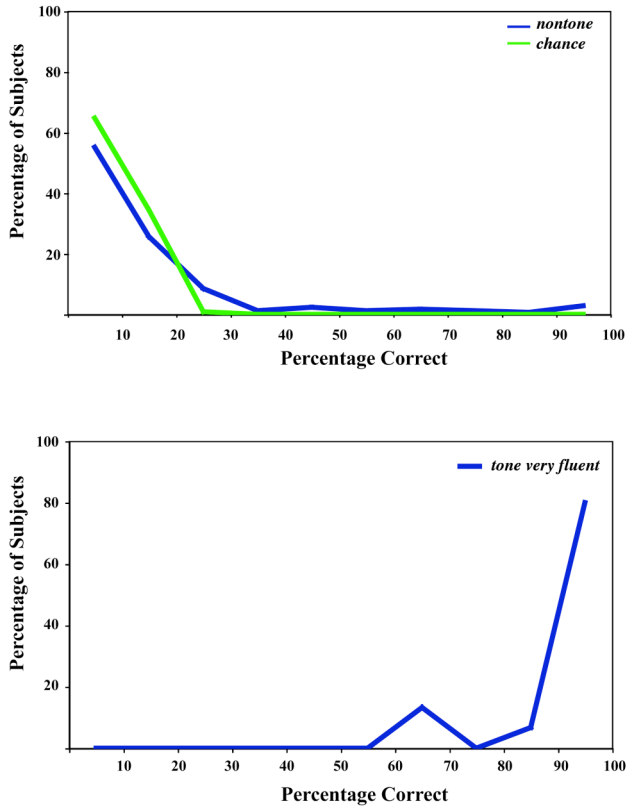

Figure 2. Upper graph: Relative distribution of scores on the test of absolute pitch among the nontone language speakers, together with the distribution expected from chance performance. Lower graph: Relative distribution of scores among those who spoke a tone language very fluently. Adapted from Deutsch et al., 2009a.

The strong relationship between the prevalence of absolute pitch on the one hand and fluency in speaking a tone language on the other is also shown in Figure 2. The upper graph shows the relative distribution of scores of all the *nontone* language speakers, together with the hypothetical distribution of scores expected from chance performance. The two plots are strikingly similar, with a very slight increase in the prevalence of absolute pitch in the 90-100% region. In contrast, the lower graph shows the relative distribution of scores in the *tone very fluent* group. As can be seen, the performance level of most of these subjects was in the 90-100% region.

The next issue we addressed concerned the country in which the subjects had received their music education. We considered only the *tone very fluent* subjects, and divided these into two groups: those who had been born in the U.S. or who had arrived in the U.S. before age 9 (*tone very fluent early arrivals*), and those who had arrived in the U.S. after age 9 (*tone very fluent late arrivals*). We also reanalyzed the data that we had earlier collected on students at the *Central Conservatory of Music* in Beijing. In order to compare this group directly with the *USC tone very fluent* subjects, we divided them into the same two subgroups by age of onset of musical training (i.e., ages 2-5, and 6-9), and we calculated the average percentage correct rather than using our earlier measure of the percentage of subjects who obtained at least 85% correct.

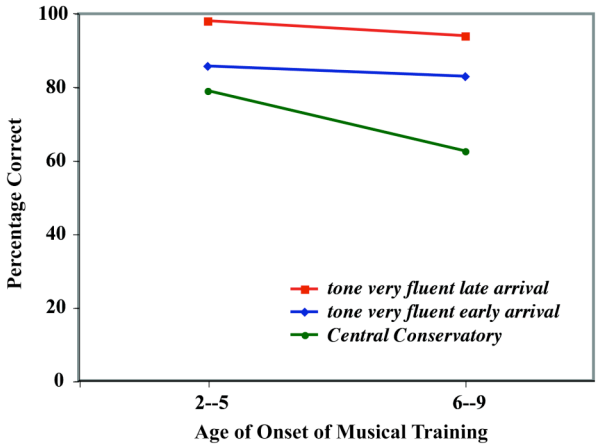

Figure 3. Percentage correct responses on the test of absolute pitch, for the three groups who spoke a tone language very fluently. See text for details. Adapted from Deutsch et al., 2009a.

The results are shown in Figure 3. The average performance of both the *USC tone very fluent* groups was higher than that of the *Central Conservatory* group. This effect was not statistically significant, but since the *Central Conservatory* group had all received their music education in China, we can conclude that the very high performance levels of the *USC tone very fluent* subjects cannot be attributed to country of music education. We can also speculate that the higher performance of the *tone very fluent USC* group compared with the *Central Conservatory* group may have been due to the *Central Conservatory* students having had extensive experience with Asian musical scales, rather than the scales of traditional Western music, and this could have interfered with our note naming task.

Finally, we can ask why some people who are not speakers of tone language acquire absolute pitch. Given our findings, those who are born into families of musicians would be at an advantage, since they would have been exposed to musical notes together with their names very early in life – often before they began musical training. As for the remainder, we may conjecture that such individuals may have a critical period of unusually long duration, so that it extends to the age at which they can begin taking music lessons.

III. THE PITCH RANGES OF SPEECH IN TWO LINGUISTIC COMMUNITIES

Our second study (Deutsch et al., 2009b) was also motivated by the findings of Deutsch et al. (2004) showing that native speakers of Mandarin and Vietnamese displayed remarkable consistency in the pitches with which they enunciated lists of words in their native languages, on two separate days. This consistency within individual subjects raises the question of whether, assuming a homogeneous linguistic community, once might find consistency across subjects also.

Earlier, Deutsch and coworkers (cf. Deutsch, 1992) had hypothesized that, through long-term exposure to speech in the environment, the individual develops a mental representation of the expected pitch range and pitch level of speech (for male and female speakers analyzed separately) in his or her linguistic community. More specifically, it was conjectured that this representation includes a delimitation of the octave band in which the largest numbers of pitch values occurs. This representation then influences both how speech is perceived, and also the listener's own speech output. It is interesting to note here that the pitch range of speech for individual speakers has frequently been shown to be roughly an octave, for both male and female speakers, and across different languages and dialects (cf. Dolson, 1994; Hudson and Holbrook, 1982; Kunzel, 1989; Xue et al, 2002). It was hypothesized that, taking two linguistic communities, each of which is homogeneous linguistically, the pitch range of speech should cluster within each community, but might differ across communities.

The present study was designed to test this hypothesis by comparing the pitch levels of female speech in two communities, each of which is quite homogenous linguistically. The subjects were located in two villages situated in a relatively remote area of China, which is considered to be stable and homogeneous in terms of ethnicity and lifestyle (Blunden and Elvin, 1998). The villages are less than 40 miles apart, though because the terrain is rugged, travel between them by automobile takes several hours. The dialects spoken by the residents of the two villages are similar to each other, being in the same family as Standard Mandarin; and communication between residents of the two villages occurs without difficulty. All subjects had learned to read and speak Standard Mandarin in school.

Thirty-three female subjects participated in the experiment. They were tested in two locations: 17 subjects (average age 33.7 years) were tested in Taouan Village, near Guandu, Zushan County in Hubei Province, and 16 subjects (average age 37.6 years) were tested in Jiuying Village, Wuxi County in the municipality of Chongqing.

Each subject was given a passage of roughly 3.25 minutes in duration to read out in Standard Mandarin, and from these readings pitch samples were taken at 5 ms intervals. The pitch values were then analyzed in two ways: First, they were allocated to semitone bins (with the bins defined by the equal-tempered scale, A = 440 Hz) and for each subject the semitone bin delimiting the octave band containing the largest number of pitch values was derived. Taking all those from Taoyuan Village, the F0 values included in the octave bands of the individual subjects comprised 98.91% of the total, and taking all those from Jiuying Village, these values comprised 97.24% of the total. This close relationship to the octave is in accordance with other literature, and with the proposed hypothesis.

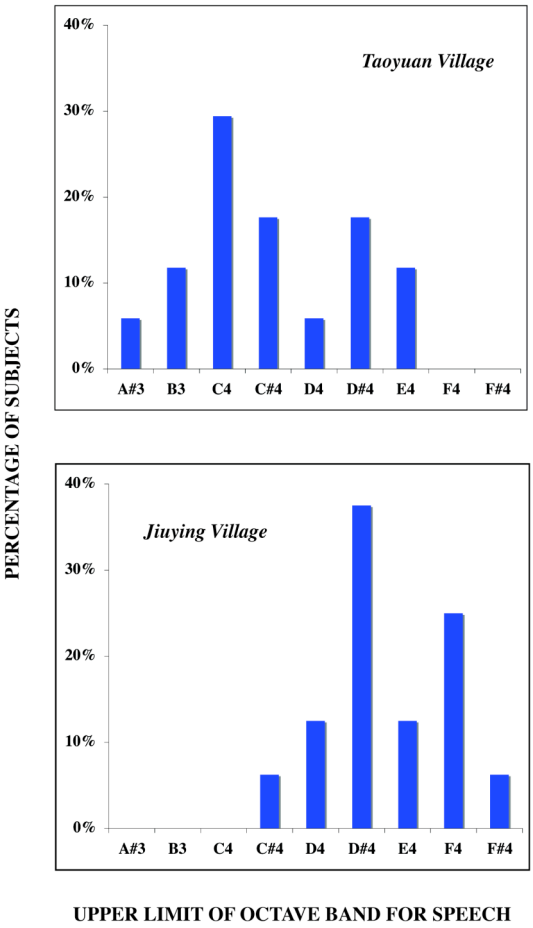

Figure 4. Upper limits of the octave band for speech in the two villages, plotted in semitone bins. See text for details. Adapted from Deutsch et al, 2009b.

The F0s were higher overall for the subjects tested in Jiuying Village compared with those in Taoyuan Village; the average F0s (i.e., averaged over a log scale) were 231.4 Hz for Jiuying Village and 200.6 Hz for Taoyuan Village. This difference between the two villages was highly significant [$F(1,31) = 19.106$; $p < 0.001$]. As a further analysis, we tested the hypothesis that the octave bands for speech of the individual subjects would cluster within each village, but would differ across villages. Figure 4 displays the percentages of subjects for whom the upper limit of the octave band fell in each semitone bin, shown separately for each village. As can be seen, the values indeed clustered within each village, and differed overall across villages by about 3 semitones. This difference between the two villages was also found to be highly significant [$F(1,31) = 19.803$; $p < 0.001$].

The findings from this experiment are in accordance with the hypothesis that the pitch of speech, at least for tone language speakers, is determined by a long term mental representation that is acquired through exposure to the speech of others. Our findings therefore provide further evidence that absolute pitch is speech-related, and that it can carry over to music under appropriate circumstances.

IV. CONCLUSION

In this paper we have presented two new studies that provide evidence for the hypothesis that absolute pitch originally evolved to subserve speech, and is still employed in the process of speech communication, at least by tone language speakers. We acknowledge that other factors also contribute to the probability of acquiring this ability. There may also be a genetic component (though so far none has been identified); age of onset of onset of musical training has been shown to be a factor (Baharloo et al, 1998; Deutsch et al., 2006); and there is also evidence that type of early musical training plays a role (particularly *fixed do* versus *moveable do*; Miyazaki, 2004). At all events, the findings described in this paper show that fluency in speaking a tone language must exert a very strong influence on the predisposition to acquire absolute pitch.

REFERENCES

Bachem, A. (1955). Absolute pitch. *Journal of the Acoustical Society of America, 27,* 1180–1185.

Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., & Freimer, N. B. (1998). Absolute pitch: An approach for identification of genetic and nongenetic components. *American Journal of Human Genetics, 62*, 224–231.

Blunden, C., & Elvin, M. (1998). *Cultural Atlas of China*, 2nd ed. (Checkmark Books, NewYork).

Deutsch, D. (1992). Some new pitch paradoxes and their implications. *Philosophical Transactions of the Royal Society of London, Series B, 336,* 391–397.

Deutsch, D. (2002). The puzzle of absolute pitch. *Current Directions in Psychological Science, 11,* 200–204.

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

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H.-S. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of the Acoustical Society of America, 119*, 719–722.

Deutsch, D., Dooley, K., Henthorn, T., & Head, B. (2009a) 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, D., Le, J., Shen, J., & Henthorn, T. (2009b). The pitch levels of speech in two Chinese villages. *JASA Express Letters*.

Dolson, M. (1994). The pitch of speech as a function of linguistic community. *Music Perception,11,* 321–331.

Hudson, A. I., & Holbrook, A. (1982). Fundamental frequency characteristics of young black adults: Spontaneous speaking and oral reading. *Journal of speech and hearing research, 25,* 25–28.

Kunzel, H. J. (1989). How well does average fundamental frequency correlate with speaker height and weight? *Phonetica 46,* 117–125.

Miyazaki, K. (2004). How well do we understand absolute pitch? *Acoustics, Science and Technology, 25,* 426–432.

Profita, J., & Bidder, T. G. (1988). Perfect pitch. *American Journal of Medical Genetics. 29,* 763–771.

Xue, S. A., Hagstrom, F., & Hao, J. (2002). Speaking fundamental frequency characteristics of young and elderly bilingual Chinese-English speakers: A functional system approach, *Asia Pacific Journal of Speech, Language, and Hearing. 7,* 55–62.