Absolute pitch among students in an American music conservatory: Association with tone language fluency

Diana Deutsch and Kevin Dooley
Department of Psychology, University of California, San Diego, La Jolla, California 92093

Trevor Henthorn
Department of Music, University of California, San Diego, La Jolla, California 92093

Brian Head
Thornton School of Music, University of Southern California, Los Angeles, California 90089

(Received 6 October 2008; revised 15 January 2009; accepted 18 January 2009)

Abstract—Absolute pitch (AP), the ability to name a musical note in the absence of a reference note, is extremely rare in the U.S. and Europe, and its genesis is unclear. The prevalence of AP was examined among students in an American music conservatory as a function of age of onset of musical training, ethnicity, and fluency in speaking a tone language. Taking those of East Asian ethnicity, the performance level on a test of AP was significantly higher among those who spoke a tone language very fluently compared with those who spoke a tone language fairly fluently and also compared with those who were not fluent in speaking a tone language. The performance level of this last group did not differ significantly from that of Caucasian students who spoke only nontone language. Early onset of musical training was associated with enhanced performance, but this did not interact with the effect of language. Further analyses showed that the results could not be explained by country of early music education. The findings support the hypothesis that the acquisition of AP by tone language speakers involves the same process as occurs in the acquisition of a second tone language. 

I. INTRODUCTION

Absolute pitch (AP)—the ability to name a musical note in the absence of a reference note—is very rare in North America and Europe, with an estimated prevalence of less than 1 in 10,000 (Bachem, 1955; Profita and Bidder, 1988). This ability is not necessarily accompanied by superior performance on other auditory or musical tasks (Bachem, 1937; Rakowski, 1972; Lockhead and Byrd, 1981; Rakowski and Morawska-Büngeler, 1987; Miyazaki, 1992, 2004; Burns and Campbell, 1994; Miyazaki and Rakowski, 2002; Fujisaki and Kashino, 2002), and its genesis is unclear.

Perhaps the most surprising feature of AP is its rarity. Most people readily identify hundreds of musical passages by name, yet AP involves choosing between only the 12 notes of the chromatic scale: C, C#, D, and so on. Furthermore, this ability is rare even among professional musicians, who spend tens of thousands of hours reading musical scores and playing the notes they read. Indeed, the vast majority of documented attempts to train adult musicians to acquire AP have been strikingly unsuccessful (Takeuchi and Hulse, 1993; Ward, 1999; Rakowski and Miyazaki, 2007).

Reports from possessors of AP indicate that the earlier the onset of musical training the higher the probability of acquiring this ability (Bachem, 1955; Miyazaki, 1988; Profita and Bidder, 1988). Large-scale studies involving both surveys (Baharloo et al., 1998) and direct tests (Deutsch et al., 2006) have found AP to be most prevalent among those who had begun musical training by ages 4–5, less prevalent between ages 6–8, and very rare after age 9. This association between the possession of AP and age of onset of musical training has led to the surmise that a critical period beginning at around age 2 or 3 is involved in its acquisition (see, for example, Takeuchi and Hulse, 1993). However since it is impractical to initiate formal musical training at an earlier age, the possibility arises that, in principle, this critical period might extend down to infancy. Indeed, Saffran and Griesentrog (2001) found that infants were able to perform a perceptual learning task that required referring to the absolute pitches of tones.

There is an intriguing parallel between the time frames involved in acquiring speech and language, on the one hand (Kuhl et al., 1992), and AP on the other. Concerning first language acquisition, there are rare documented cases of children who had been socially isolated early in life and were later placed in a normal environment; these children were found unable to acquire normal language (Lane, 1976; Curtiss, 1977). Studies of recovery of speech following brain injury have indicated the same time frame: The prognosis for recovery has been found most positive for those whose injury occurred before age 6, less positive between ages 6 and 8, and very poor after puberty (Dennis and Whitaker, 1976; Bates, 1992; Duchowny et al., 1996; Doupe and Kuhl, 1999). Second language acquisition beyond early childhood occurs far more readily than does acquisition of a first language, presumably because the basic underlying circuitry has already been established. However, acquisition of a second
language is also subject to a critical period: It occurs extremely readily in infancy (Kuhl et al., 2003), with proficiency in its acquisition declining with increasing age of initial exposure to the second language, reaching a plateau in adulthood (Lennenberg, 1967; Johnson and Newport, 1989; Newport, 1990).

Given the similarity between the time frames for acquiring speech, on one hand, and AP, on the other, it has been suggested that AP was originally packaged in with other features of speech, so that it might be expected to be heavily influenced by the speech-related critical period (Deutsch, 2002; Deutsch et al., 2004, 2006). This hypothesis was first proposed based on the finding that speakers of two tone languages (Vietnamese and Mandarin) displayed a remarkably precise and stable form of AP in reciting lists of words on different days (Deutsch et al., 2004). In tone languages, words take on arbitrarily different lexical meanings depending on the “tones” in which they are pronounced. Tones are defined by both their pitch heights and contours. In Mandarin, for example, the first tone is characterized as high and level, the second as mid-high and rising, the third as initially low and descending and then rising, and the fourth as high and falling. For example, the word “ma” means “mother” when spoken in the first tone, “hemp” in the second, “horse” in the third, and a reproach in the fourth. When, therefore, speakers of Mandarin hear the word “ma” spoken in the first tone, and attribute the meaning “mother,” or when they hear “ma” spoken in the second tone and attribute the meaning “hemp,” they are using pitch (along with other features of speech) in the process of attributing a verbal label to the sound. Analogously, when people with AP hear the tone F♯ and identify it as “F♯,” or hear the note E and identify it as “E,” they are also using pitch to attribute a verbal label to the sound.

On this line of reasoning, it was hypothesized that in cultures where tone languages are spoken, infants generally acquire AP for the tones of their language during the critical period in which they acquire other features of their native language (Deutsch, 2002). When they reach the age at which they can begin musical training, they acquire AP for musical tones in the same way as they would acquire the tones of a second tone language (see, for example, Wayland and Guion, 2004). For such individuals, therefore, the acquisition of AP for musical tones should also be subject to a critical period, but its overall prevalence should be considerably higher than for those who had instead acquired a nontone language in infancy and so would not have had the opportunity to develop the basic underlying circuitry at that time.

In a study designed to evaluate this hypothesis, two large groups of music students were administered a direct on-site test of AP, without self-selection from among the target populations. The first group comprised students at the Central Conservatory of Music in Beijing; these were all tone language speakers. The second group comprised students at Eastman School of Music who were nontone language speakers. The data from both groups displayed large and orderly effects of age of onset of musical training, and importantly, the speakers of tone language far outperformed those of nontone language for all levels of age of onset of training (Deutsch et al., 2006).

The above findings are in accordance with the hypothesis that the acquisition of AP is subject to a speech-related critical period. However, the findings are also consistent with a genetic explanation for the difference between the two groups in prevalence of AP. The present study was designed to evaluate the hypothesis that AP is influenced by a speech-based critical period, holding genetic factors constant.

We compared four groups of subjects. The first group, termed nontone, consisted of those who were Caucasian and spoke only nontone language. The remaining subjects were of East Asian ethnic heritage, with both parents speaking an East Asian tone language. These subjects were assigned to three groups in accordance with their responses to a questionnaire. Those in group tone very fluent responded that they spoke an East Asian tone language very fluently; those in group tone fairly fluent responded that they spoke an East Asian tone language fairly fluently; those in group tone non-fluent responded that they were not fluent in speaking a tone language. By comparing the prevalence of AP in these four groups, we evaluated the relative contributions of ethnic heritage and language to the probability of acquiring AP.

II. METHOD

A. Procedure

Subjects were administered a test for AP that was identical to that given in the study of Deutsch et al. (2006). They were presented successively with the 36 tones spanning three octaves from C3 (131 Hz) to B5 (988 Hz) and were asked to write down the name of each tone (C, C♯, D, and so on) 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 stimuli were piano tones that were generated on a Kurzweil K2000 synthesizer tuned to A4=440 Hz. The tones were all 500 ms in duration and were presented in three blocks of 12, with 4.25-s intervals between onsets of tones within a block and 39-s pauses between blocks. 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 blocks. The tones were presented via a compact disk player, amplifier, and two loudspeakers.

Following the test, the subjects filled out a questionnaire that enquired into their music education, ethnic heritage, where they and their parents had lived, languages they and their parents spoke, and how fluently they spoke each language. Specifically concerning linguistic fluency, the subjects chose between the following options: (a) I speak the language very fluently, (b) I speak the language fairly fluently, and (c) I can understand the language, but don’t speak it fluently.

B. Subjects

There were 203 subjects in this study. These were 110 males and 93 females, with average age 19.5 years (range
18–33 years). 162 subjects were first- or second-year students taking a required course at Thornton School of Music at the University of Southern California (USC); 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 in this experiment. All subjects were free of hearing problems, as indicated by self-report.

Based on their responses to the questionnaire, the subjects were divided into four groups. Those in group nontone were Caucasian and spoke only nontone language fluently. These were \( n = 176 \), 98 males and 78 females, with average age 19.4 years (range 18–33 years). The remaining subjects were all of East Asian (Chinese or Vietnamese) ethnic heritage, with both parents primarily speaking an East Asian tone language. These were assigned to three further groups in accordance with their responses to the questionnaire. Those in group tone very fluent reported that they spoke an East Asian tone language “very fluently.” These were \( n = 15 \), 6 males and 9 females, with average age 20.3 years (range 18–28 years). Those in group tone fairly fluent reported that they spoke a tone language “fairly fluently.” These were \( n = 7 \), 4 males and 3 females, with average age 21.3 years (range 18–29 years). Those in group tone nonfluent responded “I can understand the language, but don’t speak it fluently.” These were \( n = 5 \), 2 males and 3 females, with average age 18.6 years (range 18–19 years).

Each of the four groups was divided into subgroups based on age of onset of musical training: One subgroup consisted of those who had begun musical training at ages 2–5. For nontone, \( n = 49 \); for tone very fluent, \( n = 12 \); for tone fairly fluent, \( n = 3 \); and for tone nonfluent, \( n = 4 \). Another subgroup consisted of those who had begun musical training at ages 6–9. For nontone, \( n = 87 \); for tone very fluent, \( n = 3 \); for tone fairly fluent, \( n = 4 \); and for tone nonfluent, \( n = 1 \). A third subgroup of nontone language speakers (\( n = 40 \)) comprised those who had begun musical training at \( \geq 10 \) years. However, all the subjects of East Asian ethnic heritage had begun musical training at \( \leq 9 \). For this reason, the data from the nontone language speakers with onset of musical training at \( \geq 10 \) are displayed but not compared statistically with that of any of the other groups.

### III. RESULTS

Figure 1 shows, for each subgroup, the mean percentage correct on the test for AP. As is evident, all groups showed clear effects of age of onset of musical training: Performance levels on this test were higher for those who had begun musical training at ages 2–5 than at ages 6–9 (and also than at age \( \geq 10 \) in the case of the nontone language speakers). This association with age of onset of musical training is in accordance with earlier findings (Baharloo et al., 1998; Deutsch et al., 2006).

It can also be seen that those subjects who stated that they spoke a tone language very fluently (tone very fluent) displayed remarkably high performance levels on this test. Indeed, their performance was far higher than that of the nontone language speakers (nontone). It was also far higher than that of the subjects of same ethnicity but who did not speak a tone language fluently (tone nonfluent). In addition, it was higher than that of the tone fairly fluent speakers, which in turn was higher than that of the tone nonfluent speakers and of the nontone language speakers (nontone).

To make statistical comparison between the groups, an overall test for the effect of gender was first performed and found to be nonsignificant \( (F < 1) \); the data from males and females were therefore combined in all further analyses. A \( 4 \times 2 \) analysis of variance (ANOVA) was performed, with group (nontone, tone very fluent, tone fairly fluent, tone nonfluent) and age of onset of musical training (2–5 and 6–9) as factors. A significant effect of age of onset was found \( [F(1, 148) = 5.15, p = 0.025] \). In addition, a highly significant effect of group was found \( [F(3, 148) = 35.43, p < 0.001] \). The interaction between group and age of onset was nonsignificant \( (F < 1) \).

On post hoc comparisons, the performance level was significantly higher for tone very fluent compared with nontone \( (p < 0.001, \text{ one-tailed}) \), for tone very fluent compared with tone nonfluent \( (p < 0.001, \text{ one-tailed}) \), and for tone very fluent compared with tone fairly fluent \( (p < 0.001, \text{ one-tailed}) \). The performance level for tone fairly fluent was significantly higher than for nontone \( (p = 0.003) \) and also higher than for tone nonfluent (though the latter comparison was not statistically significant; \( p > 0.05 \)). The performance level for the tone nonfluent group did not differ significantly from that for the nontone group \( (p > 0.05) \). These findings indicate strongly that the differences in performance levels between the groups were determined by language rather than ethnicity. Indeed, taking those subjects of East Asian heritage, a regression analysis found that fluency in speaking a tone language was a highly significant predictor of performance \( (R^2_{\text{adj}} = 0.54, F(2, 24) = 16.09, p < 0.001) \).
The strong relationship between the prevalence of AP and fluency in speaking a tone language is also reflected in the data shown in Fig. 2. The upper graph shows the relative distribution of scores of all nontone language speakers (group nontone), together with the hypothetical distribution of scores expected from chance performance. Lower graph: relative distribution of scores in the group tone very fluent.

One issue that arises from these findings concerns the country in which the subjects received their music education. To evaluate this factor, comparison was made, taking only the subjects in the 90%–100% region. The lower graph shows, in contrast, 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.

Interestingly, the overall effect of group was also nonsignificant [$F(2,89)=1.76, p > 0.05$]. In post hoc comparisons, the performance of the TVF early arrivals was not significantly different from either the TVF late arrivals ($p > 0.05$) or the Central Conservatory group ($p > 0.05$). However, the performance level of the TVF late arrivals was significantly higher than that of the Central Conservatory.

Given the remarkably high performance of the two tone very fluent groups, we compared their performance levels with those obtained by Deutsch et al. (2006) from the students at the Central Conservatory of Music in Beijing who had been administered the identical test for AP. The subjects tested in Beijing had all been born and grown up in China. In this study, we had compared the numbers of subjects in each group who obtained at least 85% correct on the test. In the present experiment, as with the other analyses we performed here, we instead employed the more sensitive measure of percentage correct for each subject. We divided the subjects from the Central Conservatory of Music into the same two categories based on age of onset of musical training as for the subjects in the present experiment, i.e., ages 2–5 and ages 6–9. The Central Conservatory subgroup with age of onset of musical training 2–5 consisted of 46 subjects, and the subgroup with age of onset of musical training 6–9 consisted of 34 subjects.

Figure 3 displays the average performance levels of each of these six subgroups. A 2 x 2 ANOVA was carried out with age of onset (2–5 and 6–9) and group (TVF early arrival, TVF late arrival, and Central Conservatory) as factors. The effect of age of onset was nonsignificant ($F<1$). However, this was presumably attributable to the ceiling effect within the TVF groups described above. Taking the Central Conservatory group alone, where performance was not at ceiling, the effect of age of onset was clearly significant [$F(1,78) = 5.69, p=0.02$].

Central Conservatory
group ($p=0.013$), and the performance level of the **TVF early arrivals** was also higher as a trend than that of the **Central Conservatory** group. This pattern of results shows that the higher performance level among the **tone very fluent** USC students in this study cannot be attributed to country of early music education.

### IV. DISCUSSION

The findings presented here indicate strongly that the enhanced prevalence of AP among tone language speakers found by Deutsch et al. (2006) was language-related rather than genetic in origin. The present findings further show that when degree of fluency in speaking a tone language is held constant, differences in prevalence of AP among the various groups cannot be explained by the country in which the subjects had received their music education.

Specifically, with respect to the genetic issue, we found, among subjects of the same ethnic heritage, significant differences in performance level depending on degree of fluency in speaking a tone language. In addition, we found no significant difference in performance level depending on degree of fluency among subjects of the same ethnic heritage, significant differences in performance level between the genetically East Asian **tone nonfluent** group and the genetically Caucasian **nontone** group. With respect to the surmise that the differences we obtained might be related to country of early music education, we found no significant difference in overall performance level between the **tone very fluent (TVF)** subjects and those tested in the study of Deutsch et al. (2006) who were students at the Central Conservatory of Music in Beijing and who had all received their music education in China. Furthermore, we found no significant difference between the (USC) **TVF early arrivals** and **TVF late arrivals**, who all spoke a tone language very fluently. Indeed, the prevalence of AP was higher as a trend among the **TVF early arrivals**, who had received their music education in the U.S., than among the **Central Conservatory** students, who had received their music education in China. This difference might be accounted for by the fact that many of the **Central Conservatory** students would have had extensive experience with Asian musical scales, and this could have interfered to some extent with the note naming task employed here. It is further interesting to observe that there was a trend for the **TVF late arrivals** to perform better than the **TVF early arrivals**, and we may conjecture that this superior performance resulted from their more extensive experience with tone language. Furthermore, since the **TVF late arrivals** were studying at USC we may conjecture that they would have been more immersed in the tradition of Western tonal music, and so with Western musical scales, than were the **Central Conservatory** students; this could account for the higher performance levels of the **TVF late arrivals** compared with the **Central Conservatory** group.

It is also interesting to note that, taking the subjects of Chinese or Vietnamese ethnic heritage, those who spoke a tone language fluently were considerably more numerous than were those who were not fluent in speaking a tone language, and further that for all subjects of Chinese or Vietnamese ethnic heritage (including those born in the U.S.) both parents spoke a tone language fluently. This contrasts with the assumption that is frequently made that East Asian students in music schools in the U.S. would include many who speak only English. The present findings show that using the criterion of “ethnicity” without considering linguistic factors could lead to incorrect conclusions about the genesis of AP [see also the survey study by Gregersen et al. (2001) and the Comment on this study by Henthorn and Deutsch (2007)].

It is further interesting to note that all subjects in the **tone very fluent**, **tone fairly fluent**, and **tone nonfluent** groups would have been exposed to tone language in infancy. This shows that exposure to tone language during the critical period for language acquisition is not, in itself, sufficient to produce an advantage for the acquisition of AP. Our findings are, however, in accordance with those of others showing substantial first language attrition in children who, having initially been exposed to one language, later primarily speak a different language (see, for example, Kaufman and Aronoff, 1991; Nicoladis and Grabois, 2002; Ventureyra et al., 2004). We must conclude from our present findings that continuity in speaking a tone language is an important factor in the probability of acquiring this ability.

The present findings leave open the question of the genesis of AP among individuals who speak only nontone language. On the present line of reasoning, there would be an advantage to those who are born into families of musicians since they would have had the opportunity to associate musical tones with their names very early in life—frequently before they begin formal musical training. This could account for findings of familial aggregation of AP (Bachem, 1955; Profita and Bidder, 1988; Baharloo et al., 1998) though such findings are also in accordance with genetic explanations. It has further been conjectured (Deutsch, 2002, 2006; Deutsch et al., 2004) that those who are not born into a strongly musical environment and nevertheless acquire AP may have a language-related critical period of unusually long duration, so that it extends to the age at which they can begin taking music lessons. Such an extended critical period might well have a genetic basis. The finding of an enhanced leftward asymmetry of the planum temporale among Western AP possessors (Schlaug et al., 1995; Zatorre et al., 1998; Keenan et al., 2001) would indicate that hardwired, and so possibly genetic, factors could be involved here. Nevertheless, the factors determining the acquisition of AP by non-tone language speakers remain unresolved.

### ACKNOWLEDGMENTS

We are grateful to David Walsh, Robert A. Cutietta, Frederick Lesemann, and Sharon Lavery for their help in various phases of the experiment and to Rachael Lapidis and Jing Shen for discussions.


