Music Perception
Spring 2004, Vol. 21, No. 3, 339–356

Absolute Pitch, Speech, and Tone Language: Some Experiments and a Proposed Framework

DIANA DEUTSCH & TREVOR HENTHORN
University of California, San Diego

MARK DOLSON
Creative Advanced Technology Center, Scotts Valley, CA

Absolute pitch is generally considered to reflect a rare musical endowment; however, its characteristics are puzzling and its genesis is unclear. We describe two experiments in which native speakers of tone languages—Mandarin and Vietnamese—were found to display a remarkably precise and stable form of absolute pitch in enunciating words. We further describe a third experiment in which speakers of English displayed less stability on an analogous task. Based on these findings, and considering the related literatures on critical periods in speech development, and the neurological underpinnings of lexical tone, we propose a framework for the genesis of absolute pitch. The framework assumes that absolute pitch originally evolved as a feature of speech, analogous to other features such as vowel quality, and that speakers of tone language naturally acquire this feature during the critical period for speech acquisition. We further propose that the acquisition of absolute pitch by rare individuals who speak an intonation language may be associated with a critical period of unusually long duration, so that it encompasses the age at which the child can take music lessons. We conclude that the potential to acquire absolute pitch is universally present at birth, and that it can be realized by enabling the infant to associate pitches with verbal labels during the critical period for speech acquisition.

ABSOLUTE pitch, which is generally defined as the ability to name or produce a note of particular pitch in the absence of a reference note, is extremely rare in our culture, with an estimated prevalence of less than 1 in 10,000 in the general population (Bachem, 1955; Profita & Bidder, 1988; Takeuchi & Hulse, 1993). Because of its rarity, and because most famous

Address correspondence to Diana Deutsch, Department of Psychology, University of California, San Diego, La Jolla, CA 92093. (e-mail: ddeutsch@ucsd.edu)
ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

composers and performers—such as Bach, Beethoven, Mozart, Rubenstein, Toscanini and Heifetz—were known to possess it, absolute pitch is often regarded as a mysterious and exceptional musical endowment. However, its characteristics are puzzling and its genesis is unclear. In this article, we review some of the perplexing features of absolute pitch, propose a theoretical framework that accommodates these features, and describe a set of findings that support our proposed framework.

Although absolute pitch is particularly prevalent among highly accomplished musicians, it is not necessarily accompanied by superior performance on other musical processing tasks. For example, absolute pitch possessors often make octave errors in assigning names to notes, particularly when different musical instruments are involved, and they do not necessarily outperform others in making judgments of octave register (Bachem, 1937; Carroll, 1975; Lockhead & Byrd, 1981; Miyazaki, 1988, 1989; Rakowski & Morawska-Büngeler, 1987; Takeuchi & Hulse, 1993). Furthermore, absolute pitch possessors do not necessarily outperform others in making judgments of musical interval (Burns & Campbell, 1994; Miyazaki, 1992, 1993), or on tasks involving short-term memory for pitch for which verbal labels cannot be employed as cues (Bachem, 1954; Rakowski, 1972; Rakowski & Morawska-Büngeler, 1987; Siegel, 1974).

Most interestingly, although absolute pitch is often considered to reflect an unusually strong long-term memory for pitch, people without this faculty can be shown to have surprisingly accurate long-term pitch memories under certain conditions. The tritone paradox (Deutsch, 1986, 1991, 1992, 1997) provides an example. To produce this musical illusion, two computer-generated tones are presented in succession. The tones are related by a half-octave, and are so constructed that their pitch classes are clearly defined, but their octave placement is ambiguous. When subjects are asked to determine whether such patterns ascend or descend in pitch, their judgments generally show systematic relationships to the positions of the tones along the pitch class circle. The findings with respect to this illusion show that most people possess an implicit form of absolute pitch, in that their judgments depend in an orderly way on pitch class, even though they are unable to name the tones they are judging.

It has further been found that the way the pitch class circle is oriented with respect to height is related to the language or dialect to which the individual has been exposed (Chalikia & Leinfelt, 2000; Chalikia, Miller, & Vaid, 2001; Chalikia, Norberg, & Paterakis, 2000; Chalikia, & Vaid, 1999; Dawe, Platt, & Welsh, 1998; Deutsch, 1991; Deutsch, Henthorn, & Dolson, 2004; Giangrande, 1998; Ragozzine & Deutsch, 1994) and to the pitch range of his or her speaking voice (Deutsch, Henthorn, & Dolson, 1999, 2004; Deutsch, North, & Ray, 1990). Given these findings, Deutsch (2002) hypothesized that the partial form of absolute pitch that is reflected in judgments of the tritone paradox had originally evolved to subserve speech. This hypothesis is bolstered by findings that the pitch range of the speaking voice is related to the individual's linguistic community rather than to physiological characteristics such as his or her height, weight, chest size, length of vocal tract, and so on (Deutsch, Henthorn, & Dolson, 2004; Dolson, 1994).

There is further evidence that people who are unable to name notes that are presented in isolation nevertheless evidence a partial form of absolute pitch. Terhardt and Ward (1982) and Terhardt and Seewann (1983) found that musicians who did not possess this faculty could nevertheless judge to a large extent whether a piece of music that they knew well was played in the correct key. Later, Halpern (1989) reported that musically untrained subjects, when asked to hum the first few notes of familiar songs on different occasions, were surprisingly consistent in their choice of pitches from one occasion to the next (see also Bergeson & Trehub, 2002). In a further study, Levitin (1994) asked subjects to choose a CD that contained a popular song that they knew well, and to reproduce the song by humming, whistling, or singing. The songs had been performed by only one musical group, and so had presumably been heard in only one key. On comparing the pitches of the first notes that the subjects produced with the equivalent ones on the CD, Levitin found that, when tested with two different songs, 44% of the subjects came within 2 semitones of the correct pitch for both songs.

Taking these findings together, we can conclude that absolute pitch is a complex and baffling phenomenon. It does not appear to have an explanation purely in terms of long-term memory for pitch (though this must be part of the picture), but involves verbal labeling also. However, since the naming of notes involves choosing between only 12 possibilities—the 12 notes within the octave—we would expect the task to be an easy one. Indeed, the task should be a trivial one for trained musicians, who spend thousands of hours reading musical scores and playing the notes they read. As a related point, most people easily remember melodies by name; however, the amount of information required to name a melody is vastly greater than is required to name a single note. The real puzzle concerning absolute pitch, therefore, is not why some people possess it, but rather why it is not universal. It is as though most people have a syndrome with respect to the labeling of pitches that is equivalent to color anomia (Geschwind & Fusillo, 1966), in which the patient can recognize colors, and can discriminate between them, but cannot associate them with verbal labels (Deutsch, 1988; 2002; Levitin, 1994, 1996).

Absolute Pitch in Relation to Speech and Language

The verbal labeling of pitches necessarily involves speech and language, so in searching for a framework in which to place absolute pitch, we can consider further evidence that it is tied to linguistic processing. One body of evidence concerns the neuroanatomical correlates of this faculty. Schlaug, Jancke, Huang, and Steinmetz (1995) were the first to document that musicians with absolute pitch tend to exhibit an unusual form of brain structure. In most right-handers, the planum temporale, which is critically involved in speech processing, is larger in the left than in the right hemisphere. Schlaug et al. observed that this leftward asymmetry was greater among musicians with absolute pitch than among those who did not possess this faculty. This finding indicates that absolute pitch is subserved, at least in part, by brain regions that underly speech processing (see also Keenan, Thangaraj, Halpern, & Schlaug, 2001; Schlaug, 2003; and Zatorre, Perry, Beckett, Westbury, & Evans, 1998).

Another body of evidence concerns an intriguing parallel between the critical periods involved in the acquisition of speech and language on the one hand, and the acquisition of absolute pitch on the other. In his influential book, Lennenberg (1967) pointed out that adults and young children acquire a second language in qualitatively different ways. Following puberty, such acquisition is self-conscious and labored; and even after years of experience a second language that is acquired in adulthood is spoken with a "foreign accent," and frequently with grammatical errors. Lennenberg therefore proposed that a critical period, which extends to puberty, is involved in the acquisition of speech and language.

Lennenberg's argument has received strong support from several lines of evidence (Doupe & Kuhl, 1999; Johnson & Newport, 1989; Newport, 1990; Newport, Bavelier & Neville, 2001). For example, children who had been socially isolated early in life and were later placed in a regular environment were found to be unable to acquire normal language (Curtiss, 1977; Lane, 1976). Studies of recovery of speech following brain injury have also pointed to a critical period for speech acquisition: The prognosis for recovery has been shown to be most positive if the injury occurred before age 6, less positive if it occurred between ages 6 and 8, and very poor following puberty (Bates, 1992; Dennis & Whitaker, 1976; Duchowney et al., 1996; Varyha-Khadem, Carr, Isaacs, Brett, Adams, & Mishkin, 1997; Woods, 1983). Studies of second language acquisition have confirmed this picture. Individuals who were first exposed to a second language in infancy or early childhood were found to be most proficient in that language. Proficiency was found to decline with increasing age of initial exposure to the second language, beginning at ages 4 to 6, and continuing until adulthood, when it was found to plateau out (Johnson & Newport, 1989; Newport, 1990; Newport, Bavelier, & Neville, 2001; Oyama, 1976).

Acquisition of absolute pitch, in relation to age at onset of musical training, presents a very similar picture, and this similarity is particularly striking in terms of the time frame involved. Although some degree of absolute pitch can be acquired in adulthood, this occurs only through extensive and laborious training (Brady, 1970; Cuddy, 1968). In contrast, when young children acquire absolute pitch, they generally do so automatically and unconsciously, without specific training on pitch-naming tasks. In addition, absolute pitch that is acquired in adulthood does not have the same ease and proficiency as absolute pitch acquired early in life (Takeuchi & Hulse, 1993; Ward, 1999).

Furthermore, there is considerable evidence that the prevalence of absolute pitch is inversely related to the age at onset of musical training (Bachem, 1955; Miyazaki, 1988; Profita & Bidder, 1988; Sergeant, 1969). In a survey of musicians and music students, Baharloo, Johnston, Service, Gitschier, and Freimer (1998) found that 40% of those who began musical training before age 4 reported that they possessed this faculty, compared with 27% of those who began training between ages 4 and 6; 8% of those who began training between ages 6 and 9; 4% who began training between ages 9 and 12; and 2.7% who began training after age 12. The striking correspondence between the timetables for acquisition of absolute pitch on the one hand, and speech and language on the other, suggests that these different capacities may be subserved by a common brain mechanism. Although critical periods for the development of other functions have been documented, such as the development of ocular dominance columns in the visual cortex of cats (Hubel & Wiesel, 1970), imprinting in ducks (Hess, 1973), and the development of auditory localization in barn owls (Knudsen, 1988) no other critical periods have been described that show a similar correspondence with speech and language in terms of time frame (see also Trout, 2003).

The argument for a linkage between absolute pitch and speech becomes even stronger when we consider the evidence from tone languages, such as Mandarin, Cantonese, Thai, and Vietnamese. In these languages, words take on arbitrarily different lexical meanings depending on the *tones* in which they are enunciated. Lexical tones are defined both by their pitch heights ("registers") and also by their pitch contours. In Mandarin, for example, the word "ma" means "mother" when spoken in the first tone, "hemp" in the second tone, "horse" in the third tone, and a reproach in the fourth tone.[1] So when a speaker of Mandarin hears "ma" in the first tone, and attributes to it the meaning "mother," he or she is associating a particular pitch (or combination of pitches) with a verbal label. Analogously, when a person with absolute pitch hears the note G♯, and identifies it as "G♯," he or she is also associating a particular pitch with a verbal label.

1. At a simple level of description, in Mandarin the first tone is characterized as high and level; the second tone as mid-high and rising; the third tone as low, initially falling and then rising; and the fourth tone as high and falling.

At the neurological level, there is strong evidence that the brain structures underlying the processing of lexical tone overlap with those underlying the processing of phonemes in speech. The communication of prosody and emotion has been found to be a nondominant hemisphere function, for speakers of both tone and intonation languages (Edmondson, Chan, Seibert, & Ross, 1987; Gorelick & Ross, 1987; Hughes, Chan, & Su, 1983; Ross, 1981; Tucker, Watson, & Heilman, 1977). In contrast, the processing of lexical tone has been found to be a dominant hemisphere function. For example, identification of tones has been observed to be impaired in aphasics with left-sided brain damage who are speakers of Thai (Gandour & Dardarananda, 1983; Gandour, Ponglorpisit, Khunadorn, Dechongkit, Boongird, Boonklam, & Potisuk, 1992), Mandarin (Naeser & Chan, 1980; Packard, 1986) and East Norwegian (Moen & Sundet, 1996). In accordance with these findings on brain-damaged individuals, normal Thai speakers have been reported to exhibit a right-ear advantage in dichotic listening to Thai tones when these were presented as words, though not when the same pitch patterns were presented as hums (Van Lancker & Fromkin, 1973). A third line of evidence comes from a study on normal subjects using positron emission tomography (PET). When discriminating pitch patterns in the form of Thai words, Thai subjects showed activation of the left frontal operculum (a region near Broca's area). However, when these subjects were presented with the same Thai words which had been low-pass filtered, the same pattern of brain activation did not occur (Gandour, Wong, & Hutchins, 1998).

These three lines of evidence, taken together, indicate that when tone language speakers perceive or produce pitches or pitch contours that signal meaningful words in their language, circuitry in the dominant hemisphere is involved. Given the evidence on critical periods for speech acquisition, it appears reasonable to assume that the development of such circuitry occurs very early in life, during the period in which infants acquire other features of speech, such as vowels and consonants (Doupe & Kuhl, 1999; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Lalonde, 1988). So we can hypothesize that if pitches and pitch contours are associated with meaningful words in infancy, these are later processed by the dominant hemisphere and are associated with words and verbal labeling. However, in the absence of such an early association, this brain circuitry is much less likely to develop.

The question then arises as to which features of pitch are critical to conveying lexical meaning in tone language. If these features were purely relational, then the present discussion would be irrelevant to the genesis of absolute pitch. If, however, absolute pitch were employed to signal lexical meaning, then we would have the beginnings of an explanation as to why speakers of intonation languages, such as English, find absolute pitch so difficult to acquire in adulthood. The study reported here was carried out as a test of the hypothesis that absolute pitch is indeed treated by tone language speakers as a critical feature of speech. The hypothesis entails that tone language speakers would evidence absolute pitch in speech processing, and that the memory representations of the pitches of speech sounds would be qualitatively different for speakers of tone and intonation languages.

Experiment 1

METHOD

Subjects

Seven native speakers of Vietnamese served as subjects in this experiment and were paid for their services. These were two men and five women, and they ranged in age from 27 to 56 years (mean age, 46.3 years). They had all been born and had grown up in Vietnam, and all spoke Vietnamese as their primary language. They had been living in the United States for periods ranging from a few months to 17 years. The subjects had received minimal or no musical training.

Procedure

The subjects were tested individually in two sessions, which were held on different days. In each session, the subject was seated before a microphone and was handed a list of 10 Vietnamese words to read out, at a rate of roughly one word every 2 s. The words in the list were chosen so that they spanned the range of tones in Vietnamese speech.

Apparatus

Speech was recorded onto DAT tape at a sampling rate of 44.1 kHz, using a Nakamichi microphone and a Panasonic SV-3700 Professional Digital Audio Tape Deck. The recorded samples were then transferred to a NeXT machine, where they were stored and analyzed.

Analysis Procedure

The speech samples were recorded into computer memory at a sampling rate of 44.1 kHz. The sound files were then converted to a sampling rate of 11.025 kHz, and were low-pass filtered, with cutoff frequencies of 1300 Hz for the female speakers and 650 Hz for the male speakers. Pitch (F_0) estimates were then obtained at 5-ms intervals, using a procedure derived from Rabiner and Schaffer (1978)[2] with additional signal processing by one of us (M.D.). Then for each word, the pitch estimates were averaged along the musical scale; that is, along a log-frequency continuum, thus producing an average pitch for each word. (If a speaker produced a word that resulted in fewer than 10 pitch estimates, the word was discarded from the analysis of this speaker's readings. Five out of the 70 comparisons were thus discarded.) Then for each speaker, the difference was calculated between the average pitches for each word as it was produced on the different days, and the signed differences were averaged across the words in the list.

---

2. In the Rabiner and Schaffer (1978) algorithm, six pitch detectors operate in parallel, and a decision matrix is then used to determine the best F_0 estimate.

RESULTS

Table 1 displays the numbers of subjects whose pitch difference scores fell in each 0.25-semitone bin. As can be seen, all subjects produced pitch difference scores of less than 1.1 semitone, and two of the seven subjects produced pitch difference scores of less than 0.25 semitone. The subjects must therefore have been referring to stable and precise absolute pitch templates in enunciating the list of words.

Experiment 2

The purpose of Experiment 2 was twofold. The first was to test the generality of the findings from Vietnamese subjects to speakers of a different tone language. To this end, Mandarin was chosen as the language to study. The second purpose was to explore the extent to which the pitch differences found in Experiment 1 in enunciating the same words on different days—albeit that these differences were very small—would have reflected the limitations on the precision and stability of the subjects' absolute pitch templates, compared with other factors. To this end, a more elaborate experimental design was constructed. As before, each subject participated in two sessions, which were held on different days. However, in each session the subject read out the word list twice in succession, with the readings separated by intervals of roughly 20 s. The question addressed was whether the pitch consistency found enunciating the same word list on two occasions would be greater for readings that occurred in immediate succession, compared with readings that occurred on different days. It was reasoned that a lack of difference between these two types of comparison would provide further evidence that these speakers were invoking an absolute pitch template in enunciating the words.

TABLE 1
Vietnamese Speakers: Pitch Difference Scores Produced from Reading a List of Vietnamese Words on Different Occasions

	Difference in Semitones				
	0-0.25	0.25-0.50	0.50-0.75	0.75-1.00	1.00-1.25
Across sessions: Day 1 vs. Day 2	2	2	0	2	1

NOTE—The table displays the numbers of subjects whose pitch difference scores fell in each 0.25-semitone bin.

METHOD

Subjects

Fifteen native speakers of Mandarin served as subjects in this experiment and were paid for their services. These were seven men and eight women, and they ranged in age from 22 to 40 years (mean age, 27.2 years). They had all been born and grown up in the People's Republic of China, and all spoke Mandarin as their primary language. The subjects had been living in the United States for periods ranging from several months to 6 years. With one exception, the subjects had received minimal or no musical training.

Procedure

The procedure was essentially the same as in Experiment 1, with two exceptions. First, a list of 12 Mandarin words was used, which consisted of three words in each of the four Mandarin tones, so ordered that the same tone did not occur twice in succession.

As a second difference in procedure, each subject read out the word list twice in each session, with readings separated by a time interval of roughly 20 s. Four difference scores were then obtained: between the first readings on Day 1 and Day 2; between the second readings on Day 1 and Day 2; between the first and second readings on Day 1; and between the first and second readings on Day 2.

Apparatus and Analysis Procedure

The apparatus was identical to that used in Experiment 1. The analysis procedure was also identical, except that four difference scores were now obtained for each subject. (If a subject produced a word that resulted in fewer than 10 pitch estimates, then all four comparisons involving this word were discarded from the analysis of this subject's readings. Ten of the 180 sets of comparisons were thus discarded.)

RESULTS

Table 2 displays, for each comparison, the numbers of subjects whose pitch difference scores fell in each 0.25-semitone bin. As can be seen, remarkable consistencies were again obtained. In particular, for all comparisons, 1/3 of the subjects produced difference scores of less than 0.25 semitone. Two independent statistical analyses were performed on the difference scores, to compare the pitch consistencies on immediate repetition (*within sessions*) with those across days (*across sessions*). For the first analysis, the difference scores derived from the first and second readings on Day 1 (*within sessions*) were compared with those derived from the first readings on Day 1 and Day 2 (*across sessions*). For the second analysis, the difference scores derived from the first and second readings on Day 2 (*within sessions*) were compared with those derived from the second readings on Day 1 and Day 2 (*across sessions*). For both analyses, the difference between the *within sessions* and *across sessions* difference scores did not approach significance ($F < 1$ in both cases). This result is as expected from the hypothesis that the subjects were referring to precise and stable absolute pitch templates in enunciating the words. It also leads to the surmise that,

Table 2

Mandarin Speakers: Pitch Difference Scores Produced from Reading a List of Mandarin Words on Different Occasions

	Difference in Semitones					
	0-0.25	0.25-0.50	0.50-0.75	0.75-1.00	1.00-1.25	1.25-1.50
Across sessions						
1st reading, Day 1 vs. Day 2	5	4	4	0	1	1
2nd reading, Day 1 vs. Day 2	5	5	3	1	1	0
Within sessions						
1st vs. 2nd reading, Day 1	5	4	6	0	0	0
1st vs. 2nd reading, Day 2	6	4	3	1	1	0

Note—The table displays, for each comparison, the number of subjects whose pitch difference scores fell in each 0.25-semitone bin.

although the pitch differences found on comparing the readings from different days were remarkably small, they nevertheless underestimated the precision of the subjects' absolute pitch templates.

Experiment 3

Experiments 1 and 2 had examined the performance of tone language speakers only, and this raises the issue of whether speakers of an intonation language would display the same pitch consistency on an analogous task. Experiment 3 was designed to examine this issue. The design of this experiment was the same as for Experiment 2, except that native speakers of English were employed as subjects, and a list of 12 English words was used for the readings.

METHOD

Subjects

Fourteen native speakers of English served as subjects in the experiment and were paid for their services. They were five men and nine women, and they ranged in age from 18 to 43 years (mean age, 25.2 years). They had all been born and had grown up in the continental United States, and none had a parent, grandparent, or sibling who spoke a tone or pitch-stress language.[3] Most of the subjects had received minimal or no musical training.

Procedure

The subjects were tested using the identical procedure as used for the speakers of Mandarin.

3. In pitch-stress languages such as Japanese, the pitch patterns of certain words convey lexical meaning. It is unknown whether absolute pitch levels are important in conveying lexical information in such languages.

Apparatus and Analysis Procedure

The apparatus and analysis procedure were identical to those in Experiment 2. (As in Experiment 2, if a subject produced a word that resulted in fewer than 10 pitch estimates, then all four comparisons involving this word were discarded from the analysis of this subject's readings. Twenty-seven out of the 168 sets of comparisons were discarded on this basis.)

RESULTS

Table 3 displays, for each comparison, the numbers of subjects whose pitch-difference scores fell in each 0.25-semitone bin. Two independent statistical analyses were again performed on the difference scores, to compare the pitch consistencies on immediate repetition (*within sessions*) with those across days (*across sessions*). It was found that, in contrast with the results obtained from the Mandarin speakers, the pitch consistencies obtained *within sessions* were significantly greater than those obtained *across sessions.* The difference scores derived from the first and second readings on Day 1 (*within sessions*) were significantly lower than those derived from the first readings on Day 1 and Day 2 (*across sessions*), $F(1, 13) = 9.792$, $p < .01$. Further, the difference scores derived from the first and second readings on Day 2 (*within sessions*) were significantly lower than those derived from the second readings on Day 1 and Day 2 (*across sessions*), $F(1, 13) = 16.376$, $p = .001$. The finding that English speakers, in contrast to Mandarin speakers, showed the expected increase in pitch consistency for readings that occurred in immediate succession reinforces the conclusion that the two groups of speakers were representing the absolute pitches of the words in qualitatively different ways.

Further statistical analyses confirmed this picture. On comparing the Mandarin with the English speakers in terms of consistency *within sessions,* no significant difference was obtained ($F < 1$). In contrast, there was

TABLE 3
English Speakers: Pitch Difference Scores Produced from Reading a List of English Words on Different Occasions

	Difference in Semitones					
	0-0.25	0.25-0.50	0.50-0.75	0.75-1.00	1.00-1.25	1.25-1.50
Across sessions						
1st reading, Day 1 vs. Day 2	0	6	1	3	2	2
2nd reading, Day 1 vs. Day 2	2	3	3	2	1	3
Within sessions						
1st vs. 2nd reading, Day 1	5	8	0	0	0	1
1st vs. 2nd reading, Day 2	6	4	3	0	1	0

NOTE—The table displays, for each comparison, the numbers of subjects whose pitch difference scores fell in each 0.25-semitone bin.

a significant difference between the Mandarin and English speakers in terms of consistency *across sessions,* $F(1, 27) = 7.923$; $p < .01$). In other words, the Mandarin and English speakers showed roughly the same degree of pitch consistency in enunciating their word lists twice in immediate succession, but the Mandarin speakers were significantly more consistent than the English speakers in enunciating their word lists on different days. Thus the Mandarin and English speakers performed differently on this reading task, both qualitatively and quantitatively, with the English speakers showing less pitch consistency across days.

Discussion

The findings from these three experiments are in accordance with the hypothesis that speakers of tone language employ absolute pitch as a feature of speech and that they refer to precise and stable absolute pitch templates in enunciating words. The finding that the English speakers were significantly less consistent than the Mandarin speakers in enunciating the same words on different days, whereas these two groups of speakers showed roughly the same degree of pitch consistency on immediate repetition, indicates that these two linguistic groups were processing the absolute pitch levels of speech in qualitatively different ways. This conclusion is in accordance with the evidence from the neuropsychological literature showing that, whereas pitch patterns are processed for intonation purposes primarily by the nondominant hemisphere (Edmondson et al., 1987; Gorelick & Ross, 1987; Hughes et al., 1983; Ross, 1981; Tucker et al., 1977) the processing of lexical tone primarily involves the dominant hemisphere (Gandour & Dardarananda, 1983; Gandour et al., 1992; Naeser & Chan, 1980; Packard, 1986, Moen & Sundet, 1996; Gandour et al., 1998; Van Lancker & Fromkin, 1973). The present results strongly indicate that absolute pitch is involved in such lexical processing.

The present argument that absolute pitch is employed by tone language speakers as a feature of speech carries with it the strong hypothesis that different individuals of the same sex who speak in the same dialect (and who had been exposed to the same pattern of dialects earlier in life) should match up in terms of the absolute pitch levels with which they enunciate words. Informal work has indicated that such across-speaker consistencies do indeed exist, and a formal study is currently underway to examine this issue.

Another interesting aspect of the present findings is that the English speakers, although not as consistent as the Mandarin speakers in the pitches with which they enunciated the same words on different days, nevertheless showed considerable consistency on this measure. This finding may be related to those described in the Introduction showing that speakers of intonation languages, who do not possess absolute pitch as conventionally defined, can nevertheless be shown to possess a latent form of absolute pitch under certain conditions (Deutsch, 1986, 1991; Halpern, 1989; Levitin, 1994; Terhardt & Seewann, 1983; Terhardt & Ward, 1982). We may hypothesize that in the present study also, the English speakers evidenced a latent form of absolute pitch. We may further hypothesize that the qualitative difference exhibited by the Mandarin and English speakers in terms of consistency as a function of time reflects the fact that the Mandarin speakers had been given the opportunity to associate pitches with verbal labels in infancy. The finding by Saffran and Griepentrog (2001) that 8-month-old infants were able to perform a perceptual learning task that necessitated referring to the absolute pitches of tones provides strong evidence that the ability to acquire absolute pitch is indeed present in infancy (see also McMullen & Saffran, 2004).

The present study does not address the question of whether absolute pitch, when acquired as a feature of speech, later generalizes to music. We may surmise that absolute pitch for music is acquired by speakers of tone language as though it were a feature of a second language. If this were so, then, given the literature on second language acquisition (Johnson & Newport, 1989; Newport, 1990; Newport et al., 2001), we should expect that speakers of tone language would acquire absolute pitch for music most proficiently in early childhood, and that such proficiency would decline with increasing age at onset of musical training, and level off at around puberty. In other words, tone language speakers should also show a critical period effect for the acquisition of absolute pitch for musical tones. However, on the "second language" hypothesis, we should expect that the prevalence of absolute pitch among tone language speakers should be higher overall than for speakers of intonation language; in other words, that the function relating the prevalence of absolute pitch to the onset of musical training should be displaced upward for the tone language speakers.

Given this line of reasoning, how do we explain the rare cases of absolute pitch among individuals who had not been exposed to tone language? Deutsch (2002) hypothesized that for such individuals, the critical period for acquiring absolute pitch is of unusually long duration, so that it extends to the age at which the child can begin to take music lessons. Such a predisposition for an extended critical period could be genetically determined (Baharloo et al., 1998; Profita & Bidder, 1988) and could also be associated with an unusual form of brain organization (Keenan et al., 2001; Schlaug, 2003; Schlaug et al., 1995; Zatorre et al., 1998).

The present findings also relate to those of an interesting study by Gregersen, Kowalsky, Kohn, and Marvin (1999), who conducted a survey of the prevalence of absolute pitch on students in U.S. music conservatories and in university and college music programs. A higher prevalence of absolute pitch was reported among those students who described their ethnic background as Asian than among the other students. It is difficult to make direct comparison, however, between the findings of Gregersen et al. and those of the present study, since the subjects in the former study were all musically trained, and the ages at which they began musical training were not given. Furthermore, the subjects' linguistic backgrounds in the former study were not given, and we may surmise that an unknown number of the Asian students would have been exposed to tone language, or to pitch-stress language, at least early in life. To our knowledge, no study that controls for linguistic factors has documented a higher prevalence of absolute pitch among individuals of Asian origin.

Finally, the present work has implications for the issues of modularity in the processing of speech and music (McMullen & Saffran, 2004; Patel, 2003; Peretz & Coltheart, 2003; Peretz, Gagnon, Hébert, & Macoir, 2004) and of the evolutionary origins of these two forms of communication (Wallin, Merker, & Brown, 2000). It is reasonable to assume that speech and music are each subserved by a large number of modules, some of which subserve one of these faculties only, while others subserve both faculties.

Our proposed framework suggests that the brain circuitry underlying absolute pitch originally evolved to subserve speech, and that it is now involved in processing absolute pitch for both speech and music. Reasoning along these lines, we can surmise that the capacity to process certain other features of music also may have evolved to subserve speech; taking as one example the processing of rapid frequency transitions, for which common circuitry may well be involved for both domains (Samson & Ehrle, 2003). A number of principles of auditory processing apply to both speech and music (Bregman, 1990; Deutsch, 1999; Deutsch & Feroe, 1981; Lerdahl & Jackendoff, 1983), and we may assume that common brain circuitries are involved here also. On the other hand, we can also assume that certain capacities are unique to speech alone, and certain others to music alone. For example, the communication of lexical meaning is clearly unique to language (except for the special case of "program music"). We may further surmise that the processing of harmonic relationships originally evolved to subserve music, particularly choral singing, which has been argued to have a strong evolutionary value (Wallin et al., 2000). The view taken here, then, is that there is no single solution to the question of modularity; but rather that some aspects of speech and music involve circuitry that is unique to one domain or to the other, but that other aspects of speech and music, including absolute pitch, involve circuitry that is common to both domains.[4]

4. The findings from Experiments 1 and 2 in this article were first presented at the 138th meeting of the Acoustical Society of America (Deutsch, Henthorn, & Dolson, 1999). We are grateful to Vincent Hsieh, San Hsieh, Leonard Zhang, Karin Liu, Quyen Doan, Lynsey Doan, Phi Nguyen, and Larry McClure for their contributions in different phases of the study, and to Elizabeth Bates and Elissa Newport for valuable discussions concerning critical periods for language development.

References

Bachem, A. (1937). Various types of absolute pitch. *Journal of the Acoustical Society of America, 9,* 146–151.

Bachem, A. (1954). Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, *26,* 751–753.

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.

Bates, E. Language development. (1992). *Current Opinion in Neurobiology, 2,* 180–185.

Bergeson, T., & Trehub, S. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science, 13,* 17–26.

Brady, P. T. Fixed-scale mechanism of absolute pitch. (1970). *Journal of the Acoustical Society of America*, *48,* 883–887.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Burns, E. M., & Campbell, S. L. (1994). Frequency and frequency-ratio resolution by possessors of absolute and relative pitch: Examples of categorical perception? *Journal of the Acoustical Society of America*, *96,* 2704–2719.

Carroll, J. B. (1975). Speed and accuracy of absolute pitch judgments: some latter-day results. *Educational Testing Service Research Bulletin* (RB –75-35). Princeton, NJ: Educational Testing Service.

Chalikia, M. H., & Leinfelt, F. (2000). Listeners in Sweden perceive tritone stimuli in a manner different from that of Americans and similar to that of British listeners [Abstract]. *Journal of the Acoustical Society of America*, *108,* 2572.

Chalikia, M. H., Miller, K. J., & Vaid, J. (2001). *The tritone paradox is perceived differently by Koreans and Americans.* Paper presented at the 101st Annual Convention of the American Psychological Association, San Francisco, CA.

Chalikia, M. H., Norberg, A. M., & Paterakis, L. (2000). Greek bilingual listeners perceive the tritone stimuli differently from speakers of English [Abstract]. *Journal of the Acoustical Society of America, 108,* 2572.

Chalikia, M. H., & Vaid, J. (1999). Perception of the tritone paradox by listeners in Texas: A re-examination of envelope effects [Abstract]. *Journal of the Acoustical Society of America, 106,* 2572.

Cuddy, L.L. (1968). Practice effects in the absolute judgment of pitch. *Journal of the Acoustical Society of America, 43,* 1069–1076.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern day "wild child."* New York: Academic Press.

Dawe, L. A., Platt, J. R., & Welsh, E. (1998). Spectral motion after-effects and the tritone paradox among Canadian subjects. *Perception and Psychophysics*, *60,* 209–220.

Dennis, M., & Whitaker, H. A. (1976). Language acquisition following hemidecortication: linguistic superiority of the left over the right hemisphere. *Brain and Language, 3,* 404–433.

Deutsch, D. (1986). A musical paradox. *Music Perception, 3,* 275–280.

Deutsch, D. (1988). Pitch class and perceived height: Some paradoxes and their implications. In E. Narmour & R. Solie (Eds.), *Explorations in music, the arts, and ideas: Essays in honor of Leonard B. Meyer* (pp. 261–294). Stuyvesant: Pendragon Press.

Deutsch, D. (1991). The tritone paradox: An influence of language on music perception. *Music Perception*, *8,* 335–347.

Deutsch, D. (1992). Paradoxes of musical pitch. *Scientific American*, *267,* 88–95.

Deutsch, D. (1997). The tritone paradox: A link between music and speech. *Current Directions in Psychological Science, 6,* 174–180.

Deutsch, D. (1999). Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., 349–411). San Diego: Academic Press.

Deutsch, D. (2002). The puzzle of absolute pitch. *Current Directions in Psychological Science, 11,* 200–204.
Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, *88,* 503–522.
Deutsch, D., Henthorn, T., & Dolson, M. (1999). Absolute pitch is demonstrated in speakers of tone languages. *Journal of the Acoustical Society of America*, *106,* 2267.
Deutsch, D., Henthorn, T., & Dolson, M. (2004). Speech patterns heard early in life influence later perception of the tritone paradox. *Music Perception, 21,* 357–372.
Deutsch, D., North, T., & Ray, L. (1990). The tritone paradox: Correlate with the listener's vocal range for speech. *Music Perception, 7,* 371–384.
Dolson, M. (1994). The pitch of speech as a function of linguistic community. *Music Perception, 11,* 321–331.
Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22,* 567–631.
Duchowny, M., Jayakar, P., Harvey, A. S., Resnick, T., Alvarez, L., Dean, P., et al. (1996). Language cortex representation: Effects of developmental versus acquired pathology. *Annals of Neurology, 40,* 31–38.
Edmondson, J. A., Chan, J.–L., Seibert, G. B., & Ross, E. D. (1987). The effect of right-brain damage on acoustical measures of affective prosody in Taiwanese patients. *Journal of Phonetics*, *15,* 219–233.
Gandour , J., & Dardarananda,R. (1983). Identification of tonal contrasts in Thai aphasic patients. *Brain and Language, 18,* 98–114.
Gandour, J., Ponglorpisit, S., Khunadorn, F., Dechongkit, S., Boongird, P., Boonklam, R., & Potisuk, S. (1992). Lexical tones in Thai after unilateral brain damage. *Brain and Language, 43,* 275–307.
Gandour, J., Wong, D., & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport, 9,* 2115–2119.
Geschwind, N., & Fusillo, M. (1966). Color-naming defects in association with alexia. *Archives of Neurology, 15,* 137–146.
Giangrande, J. (1998). The tritone paradox: Effects of pitch class and position of the spectral envelope. *Music Perception, 15,* 253–264.
Gorelick, P. B., & Ross, E. D. (1987). The aprosodias: Further functional-anatomic evidence for organization of affective language in the right hemisphere. *Journal of Neurology, Neurosurgery, and Psychiatry, 50,* 553–560.
Gregersen, P. K., Kowalksy, E., Kohn, N., & Marvin, E. W. (1999). Absolute pitch: Prevalence, ethnic variation, and estimation of the genetic component. *American Journal of Human Genetics, 65*, 911–913.
Halpern A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition, 17,* 572–581.
Hess, E. H. (1973). *Imprinting: Early experience and the developmental psychobiology of attachment.* New York: Van Nordstrand Reinhold.
Hubel, D. H., & Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology*, *206,* 419–436.
Hughes, C. P., Chan, J. L., & Su, M. S. (1983). Aprosodia in Chinese patients with right cerebral hemisphere lesions. *Archives of Neurology*, *40,* 732–736.
Johnson, J. S., & Newport, E. L. (1989). Critical periods in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60–99.
Keenan, J. P., Thangaraj, V., Halpern, A., & Schlaug, G. (2001). Planum temporale and absolute pitch. *Neuroimage*, *14,* 1402–1408.
Knudsen, E. I. (1988). Sensitive and critical periods in the development of sound localization. In S. S. Easter, K. F. Barald, & B. M. Carlson (Eds.) *From message to mind: Directions in developmental neurobiology*. Sunderland, MA: Sinauer Associates.
Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255,* 606–608.
Lane, H. L. (1976). *The wild boy of Aveyron*. Cambridge, MA: Harvard University Press.
Lennenberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge: MIT Press.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence for the production of learned melodies. *Perception and Psychophysics, 56,* 414–423.

Levitin, D. J. (1996). Mechanisms of memory for musical attributes. *Dissertation Abstracts International, 57*(07B), 4755. (University Microfilms No. AAG9638097).

Lockhead, G. R., & Byrd, R. (1981). Practically perfect pitch. *Journal of the Acoustical Society of America, 70,* 387–389.

McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception, 21,* 000–000.

Miyazaki K. (1989). Absolute pitch identification: Effects of timbre and pitch region. *Music Perception, 7,* 1–14.

Miyazaki, K. (1993). Absolute pitch as an inability: Identification of musical intervals in a tonal context. *Music Perception, 11,* 55–72.

Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Perception & Psychophysics, 44,* 501–512.

Miyazaki, K. (1992). Perception of musical intervals by absolute pitch possessors. *Music Perception, 9,* 413–426.

Moen, I., & Sundet, K. (1996). Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage. *Brain and Language, 53,* 267–281.

Naeser, M. A., & Chan, S. W.-C. (1980). Case study of a Chinese aphasic with the Boston diagnostic aphasia exam. *Neuropsychologia, 18,* 389–410.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14,* 11–28.

Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Critical thinking about critical periods. In E. Dupoux (Ed.) *Language, brain, and cognitive development: Essays in honor of Jacques Mehler*. Cambridge, MA: MIT Press.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5,* 261–283.

Packard, J. L. (1986). Tone production deficits in nonfluent aphasic Chinese speech. *Brain and Language, 29,* 212–223.

Patel, A. D. (2003). Language, music, syntax, and the brain. *Nature Neuroscience, 6,* 674–681.

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

Peretz, I., Gagnon, L., Hébert, S., & Macoir, J. (2004). Singing in the brain: Insights from cognitive neuropsychology. *Music Perception, 21,* 373-390.

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

Rabiner, L. R., & Schaffer, R. (1978). *Digital processing of speech signals.* Englewood Cliffs, NJ: Prentice Hall.

Ragozzine, R., & Deutsch, D. (1994). A regional difference in perception of the tritone paradox within the United States. *Music Perception, 12,* 213–225.

Rakowski, A. (1972). Direct comparison of absolute and relative pitch. In F. A. Bilsen (Ed.), *Symposium on hearing theory* (pp. 105–108). Eindhoven, The Netherlands: Instituut voor Perceptie Underzoek.

Rakowski, A., & Morawska-Büngeler, M. (1987). In search of the criteria for absolute pitch. *Archives of Acoustics, 12,* 75–87.

Ross, E. D. (1981). The aprosodias: functional –anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology, 38,* 561–569.

Saffran J. R., & Griepentrog G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology, 37,* 74–85.

Samson, S., & Ehrle, N. (2003). Cerebral substrates for musical temporal processes. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 204–216). Oxford: Oxford University Press.

Schlaug G. (2003). The brain of musicians. In I. Peretz & R. Zatorre (Eds.), The cognitive neuroscience of music (pp. 366–381). Oxford: Oxford University Press.

Schlaug, G., Jaencke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science, 267,* 699–701.

Sergeant D. (1969). Experimental investigation of absolute pitch. *Journal of Research in Music Education, 17,* 135–143.

Siegel, J. A. (1974). Sensory and verbal coding strategies in subjects with absolute pitch. *Journal of Experimental Psychology, 103,* 37–44.

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin, 113,* 345–361.

Terhardt, E., & Seewann, M. (1983). Aural key identification and its relationship to absolute pitch. *Music Perception, 1,* 63–83.

Terhardt, E., & Ward, W. D. (1982). Recognition of musical key: Exploratory study. *Journal of the Acoustical Society of America, 72,* 26–33.

Trout, J. D. (2003). Biological specializations for speech: What can animals tell us? *Current Directions in Psychological Science, 12,* 155–159.

Tucker, D. M., Watson, R. T., & Heilman, K. M. (1977). Discrimination and evocation of affectively intoned speech in patients with right parietal disease. *Neurology, 27,* 947–950.

Van Lancker, D., & Fromkin, V. (1973). Hemispheric specialization for pitch and "tone": Evidence from Thai. *Journal of Phonetics, 1,* 101–109.

Varyha-Khadem, F., Carr, L. J., Isaacs, E., Brett, E., Adams, C., & Mishkin, M. (1997). Onset of speech after left hemispherectomy in a nine year old boy. *Brain, 120,* 159–182.

Wallin, N. L., Merker, B., & Brown, S. (Eds.). (2000). *The origins of music,* Cambridge, MA: MIT Press.

Ward, W. D. (1999). Absolute pitch. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 265–298). San Diego, CA: Academic Press.

Werker, J., & Lalonde, C. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24,* 672–683.

Woods, B. T. (1983). Is the left hemisphere specialized for language at birth? *Trends in Neuroscience, 6,* 115–117.

Zatorre R. J., Perry, D. W., Beckett, C. A,. Westbury, C. F., & Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proceedings of the National Academy of Sciences, 95,* 3172–3177.