Music Perception
Spring 2004, Vol. 21, No. 3, 357–372

Speech Patterns Heard Early in Life Influence Later Perception of the Tritone Paradox

DIANA DEUTSCH & TREVOR HENTHORN
University of California, San Diego

MARK DOLSON
Creative Advanced Technology Center, Scotts Valley, CA

The tritone paradox occurs when two tones that are related by a half-octave (or tritone) are presented in succession and the tones are constructed in such a way that their pitch classes (C, C♯, D, etc.) are clearly defined but their octave placement is ambiguous. Previous studies have shown that there are large individual differences in how such tone pairs are perceived, and these differences correlate with the listener's language or dialect. We here present findings showing that perception of the tritone paradox can be heavily influenced by speech heard early in life, even for listeners who do not speak their first language fluently. Our findings point to a specific linkage between speech and music, and they also shed light on the issue of critical periods for the acquisition of intonational properties of speech.

THERE is a long history of speculation concerning relationships between speech and music. In the nineteenth century, Spencer (1857) and Darwin (1871) both conjectured that these two forms of communication have a common origin—a hypothesis that has been discussed in detail by Wallin, Merker, and Brown (2000). Findings by neuroscientists have indicated that certain elements of speech and music are subserved by overlapping neural circuitry (Besson & Schon, 2003; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Samson & Ehrle, 2003). At the perceptual level, it has been shown that infants respond strongly to musical characteristics of their mothers' speech—particularly melodic contour (Fernald, 1992)—and that pitch and durational characteristics of speech play an important role in language comprehension (Cutler, 1991; Cutler, Dahan, & van Donselaar, 1997).

Address correspondence to Diana Deutsch, Department of Psychology, University of California, San Diego, La Jolla, CA 92093. (email: 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.

The tritone paradox (Deutsch, 1986, 1991, 1992, 1997) provides further evidence for an association between speech and music, since the way this musical illusion is perceived correlates with the language or dialect that is spoken by the listener (as described later). In the study presented here, we explored the basis for this association by comparing the perceptions of listeners whose first language was Vietnamese with those of monolingual speakers of English. In our first experiment, we studied three groups of listeners: Those in the first group ("Vietnamese Late Arrival") had arrived in the United States as adults, and they spoke perfect Vietnamese but little English. Those in the second group ("Vietnamese Early Arrival") had arrived in the United States as infants or children, and they spoke perfect English but were not fluent speakers of Vietnamese. Those in the third group ("Californian English") were monolingual speakers of Californian English, whose parents were also monolingual speakers of Californian English. We found that the two Vietnamese groups were closely similar to each other in their perceptions of the tritone paradox, and that both groups differed significantly from the Californian English group.

In our second experiment we found, for Vietnamese speakers, a correlate between the way the tritone paradox was perceived and the pitch range of the listener's speaking voice. The two sets of findings, taken together, lead us to conclude that the pitch of speech heard early in life has a strong influence on perception of this musical illusion in adulthood, even for listeners who do not speak their first language fluently. Beyond pointing to a specific linkage between speech and music, our findings contribute to the issue of critical periods for the acquisition of intonational characteristics of speech, and the malleability of the perceptual system beyond such critical periods.

The tritone paradox is produced by tones that are so constructed that their note names, or pitch classes (C, C♯, D, and so on) are clearly defined, but their octave placement is ambiguous.[1] So, for example, one tone might clearly be a C, but in principle it could be middle C, or the C an octave above, or the C an octave below (see Shepard, 1964, and Risset, 1971, for descriptions of other perceptual properties of tones from which the present ones were derived.)

To produce the tritone paradox, two of these ambiguous tones are presented in succession, and these are related by a half-octave (or tritone). For example, D might be presented followed by G♯, or A followed by D♯; and so on. Most listeners, when asked to determine whether such tone pairs

1. The tones employed to generate the tritone paradox each consisted of six sinusoidal components that were separated by octaves, and their amplitudes were scaled by a bell-shaped spectral envelope (see Deutsch, 1991, for details). Because the other harmonics that usually contribute to the perceived height of a tone are missing (Houtsma, 1995; Terhardt, 1974) these tones are well defined in terms of pitch class but poorly defined in terms of height.

form ascending or descending patterns, make judgments that display systematic relationships to the positions of the tones along the pitch class circle (Figure 1): Tones in one region of the circle are perceived as higher, while tones in the opposite region are perceived as lower. Yet listeners disagree substantially as to which tone pairs are heard as ascending and which as descending, and so as to which pitch classes are heard as higher and which as lower (see, for example, Deutsch, 1986, 1991, 1992, 1997).

Deutsch, North, and Ray (1990) hypothesized that the tritone paradox, and the individual differences in the way it is perceived, might be related to the processing of speech sounds (see also Terhardt, 1974, 1991). More specifically, we hypothesized that, through extensive exposure to such sounds, the individual acquires a long-term representation of the pitch range of his or her speaking voice, and included in this representation is a delimitation of the octave band in which the largest proportion of pitch values occurs. We further hypothesized that the pitch classes delimiting this octave band for speech are taken by the listener as defining the highest position along the pitch class circle, and that this is reflected in his or her judgments of the tritone paradox. In accordance with this conjecture, Deutsch et al. (1990) found a strong correspondence between the pitch range of a subject's spontaneous speech and the way he or she perceived this musical illusion.

Deutsch (1991) further conjectured that this acquired representation of the pitch class circle is derived from exposure to speech produced by others, and that it is used both to evaluate perceived speech and to constrain

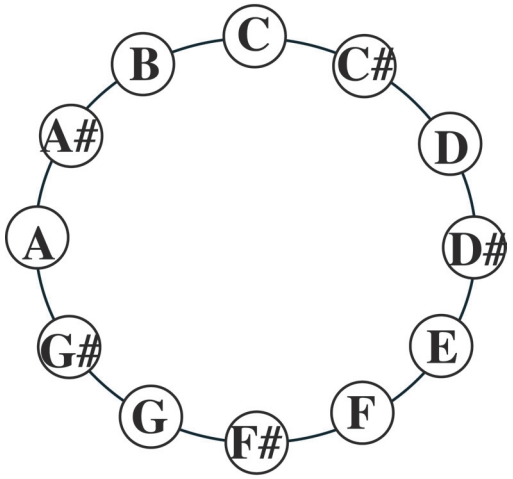

Fig. 1. The pitch class circle. The musical scale is formed by dividing the octave into 12 semitone steps, and each tone is given a name (C, C♯, D, and so on). To produce the entire scale, this succession of note names, or pitch classes, is repeatedly presented across octaves.

the listener's own speech output. On this hypothesis, the characteristics of such a template should vary among people who speak in different languages or dialects, along with other speech characteristics, such as vowel quality.

Dolson (1994) has reviewed the evidence concerning the pitch of speech in different linguistic communities, in relation to the hypothesis of an acquired pitch class template. He concluded that several lines of evidence support this hypothesis. For example, in a community of similarly aged speakers (either male or female) there is typically little variation in the pitch range of speech, as characterized by average speaker fundamental frequency. In contrast, there can be sizeable differences in pitch ranges of speech across different linguistic communities (see, for example, Majewski, Hollien, & Zalewski, 1972; Hanley, Snidicor, & Ringel, 1966; and Yamazawa & Hollien, 1992). These group differences cannot be explained on the basis of physiological characteristics of the speakers, since a number of studies have shown that speaker fundamental frequency does not correlate with measures such as laryngeal size, chest size, overall height, overall weight, and so on (Hollien & Jackson, 1973; Kunzel, 1989; Majewski et al., 1972). Furthermore, the pitch range of speech for an individual speaker generally spans roughly an octave (see, e.g., Hudson & Holbrook, 1982, and Figure 5 of the present article) and that, within a given linguistic community, the difference in pitch range between male and female speakers is generally close to an octave—roughly 10 or 11 semitones (Boe & Rakolofiringa, 1975; Hudson & Holbrook, 1981; Kunzel, 1989). This body of research therefore supports the hypothesis that the pitch range of an individual's speaking voice is determined by a pitch class template that is acquired through exposure to speech produced by others.

In a test of this hypothesis, Deutsch (1991) studied perception of the tritone paradox in two groups of subjects. Those in the first group had been born and had grown up in California, whereas those in the second group had been born and had grown up in the south of England. Both groups were tested under identical conditions; however, striking differences between them emerged. Statistically, when the Californian group tended to hear a tone pair as ascending, the southern English group tended to hear it as descending, and vice versa. This finding, in conjunction with that of Deutsch et al. (1990), indicates that the same, culturally acquired, pitch class template influences both the individual's speech production and also how he or she perceives the tritone paradox.

Studies in other laboratories have provided further evidence that differences in perception of this musical illusion are related to the languages or dialects to which listeners have been exposed. Giangrande (1998) reported that the perceptions of listeners in Boca Raton, Florida, were similar to those of the Californians studied by Deutsch (1991). In contrast, Dawe, Platt, and Welsh (1998) found that the perceptions of listeners in Hamilton, Ontario, were similar to those of listeners studied by Deutsch (1991) from

the South of England. Chalikia and colleagues have reported a number of linguistic correlates: The perceptions of English speakers in Texas differed from those of English speaking Californians (Chalikia & Vaid, 1999) and were substantially different from those of Korean speakers (Chalikia, Miller, & Vaid, 2001). The perceptions of English speakers who were raised in the upper Midwest differed from the perceptions of English speakers raised in Texas (Haugen & Chalikia, 2000), and the perceptions of Swedish listeners in Stockholm differed from those of English-speaking Texans and Californians, but were not significantly different from those of English speakers from the south of England (Chalikia & Leinfelt, 2000). The perceptions of Greek listeners in Greece differed from those of English speakers in Texas, in California, and in the south of England (Chalikia, Norberg, & Paterakis, 2000).

Assuming that perception of the tritone paradox is determined by a learned, speech-related template, the question arises as to when in life this template develops. Ragozzine and Deutsch (1994) found that, among a group of subjects who had grown up in the area of Youngstown, Ohio, statistical differences emerged between those whose parents had also grown up in this region and those whose parents had grown up elsewhere in the United States. Since parents have a particularly strong influence on speech development, these findings provided evidence that an individual's pitch class template is formed in childhood.

In a further experiment, Deutsch (1996) studied a group of subjects together with their mothers. The subjects were all Californian; however, their mothers had grown up in a number of different geographic regions. As expected, there were considerable individual differences among the mothers in the way they perceived the tritone paradox. And although all the subjects were Californian, their perceptions corresponded closely to those of their mothers, and so also differed considerably from each other. These findings therefore provided further evidence that the pitch class template that influences perception of the tritone paradox is acquired early in life. However, most of the subjects in this study were children, so it was left undetermined how well such a childhood template survives into adulthood. The present study was designed to address this issue.

Experiment 1

METHOD

Subjects

Three groups of subjects took part in the experiment and were paid for their participation. Group "Vietnamese Late Arrival" (N = 16, 6 men and 10 women; mean age, 54 years) were from South or Central Vietnam[2] and had arrived in the United States as adults. They

2. One subject was born in North Vietnam and moved to South Vietnam at age 7.

all spoke primarily Vietnamese, and none were fluent speakers of English. Group "Vietnamese Early Arrival" (N = 16 , 3 men and 13 women, mean age, 22.1 years) were university students who were also from South or Central Vietnam and had arrived in the United States as infants or young children. Although Vietnamese had been their first language, they all now spoke primarily English, and were not fluent speakers of Vietnamese.[3] Group "Californian English" (N = 10, 6 men and 4 women, mean age, 22.0 years) were university students who were monolingual speakers of English. They had been born and had grown up in California, and both their parents were also monolingual speakers of English who had grown up in California. None of the subjects had any known hearing deficits.

Stimuli and Apparatus

Each tone consisted of six octave-related sinusoids whose amplitudes were scaled by a bell-shaped spectral envelope (see Deutsch, 1991, for details). Tone pairs were created under envelopes that were placed at four different positions along the spectrum, which were spaced at half-octave intervals. Specifically, the envelopes were centered at C_4 (262 Hz), $F\sharp_4$ (370 Hz), C_5 (523) Hz), and $F\sharp_5$ (740 Hz).[4] Twelve tone pairs were generated under each of the four envelopes; these corresponded to the pitch-class pairings C-F♯, C♯-G, D-G♯, D♯-A, E-A♯, F-B, F♯-C, G-C♯, G♯-D, A-D♯, A♯-E, and B-F. The tone pairs were presented in blocks of 12; each block consisted of tones that were generated under one of the envelopes and contained one example of each of the twelve pitch class pairings. Within blocks, the tone pairs were presented in any of four orderings. The orderings were random with the restriction that the same pitch classes did not occur in any consecutive trials. Sixteen blocks were created altogether, with each of the four pitch class orderings employed once for each of the four positions of the spectral envelope. The blocks were presented in haphazard order, except that no consecutive blocks employed the same spectral envelope or the same pitch class ordering. All tones were 500 ms in duration, and there were no pauses between tones within a pair. The tone pairs were separated by 4-s intertrial intervals, and blocks were separated by 1-min pauses.

The tones were generated on a NeXT computer, recorded onto compact disc (Deutsch, 1995), and played back on a CD player. The output was passed through an amplifier and presented to subjects binaurally through headphones at a level of approximately 72 dB SPL.

Procedure

Subjects were tested in soundproof booths. On each trial a tone pair was presented, and the subject judged whether it formed an ascending or a descending pattern. All sixteen blocks were presented in each session, with a 5-min break between the eighth and the ninth blocks. Each subject served in two sessions, which were held on different days. At the beginning of each session, several practice trials were administered without feedback.

RESULTS AND DISCUSSION

For each subject, the percentage of judgments that a tone pair formed a descending pattern was plotted as a function of the pitch class of the first

3. Although the subjects in the Vietnamese Early Arrival group spoke English without a discernable Vietnamese accent, and asserted that they did not speak Vietnamese well, they were living either with their parents or with Vietnamese roommates, and so were still frequently exposed to Vietnamese speech.

4. Averaging the data from envelopes that were spaced at half-octave intervals balanced out possible effects of the relative amplitudes of the components of the tones.

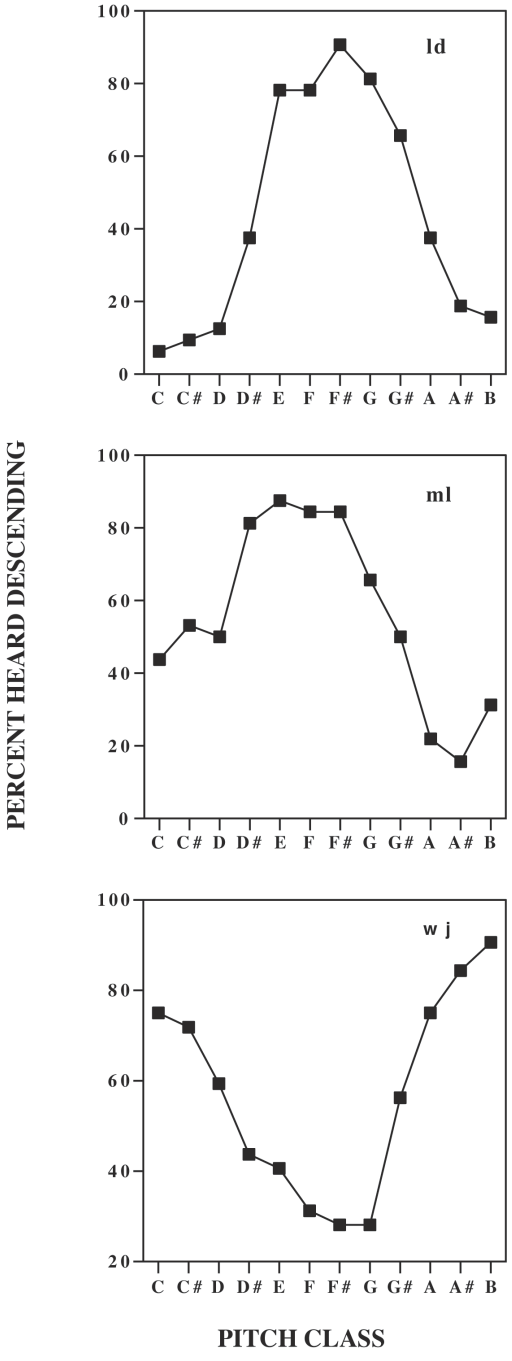

Fig. 2. Percentages of judgments that a tone pair formed a descending pattern, plotted as a function of the pitch class of the first tone of the pair. The plots from three subjects are here shown; the first from the Vietnamese Early Arrival group, the second from the Vietnamese Late Arrival group, and the third from the Californian English group.

tone of the pair. The graphs in Figure 2 show, as examples, the plots obtained from three subjects, one from each group, in each case averaged over the four spectral envelopes and over both sessions. As illustrated in these plots, the subjects' judgments were strongly influenced by the positions of the tones along the pitch class circle; however, the direction of this influence varied substantially across subjects.

In order to explore the relationship between pitch class and perceived height in each subject's judgments, we bisected the pitch class circle so as to maximize the difference between the subject's averaged scores for the two halves. (This followed the original procedure of Deutsch, Kuyper, & Fisher, 1987). The circle was then oriented so that the line of bisection was horizontal, and the data were retabulated, with the leftmost pitch class of the upper half taking the first position, its clockwise neighbor taking the second position, and so on. We then defined the pitch classes that stood at the peak of the normalized circle (Positions 3 and 4) as the subject's *peak pitch classes*.

In order to determine whether there was an overall difference in the distribution of peak pitch classes depending on the individual's native language, we compared the distributions for the two Vietnamese groups combined with that of the Californian English group. As shown in Figure 3, these distributions clearly differed from each other: For the Vietnamese group, D♯, E, F, and F♯ occurred most frequently as peak pitch classes; however, for the Californian group, C♯, D, D♯, and E occurred most frequently instead.

We next compared the two groups statistically, with respect to the overall distributions of peak pitch classes. Following the earlier procedure of Deutsch (1991), we designated those subjects whose peak pitch classes were in the range moving clockwise from A♯-B to D♯-E as in the "Californian range" and those whose peak pitch classes were in the range moving clockwise from E-F to A-A♯ as in the "opposite range." We then tabulated, for each group, the numbers of subjects whose peak pitch classes fell into these two ranges. The overall difference between the Vietnamese and Californian English groups was highly significant ($p < .001$ on a Fisher exact probability test).

Further analysis showed that the difference between the Vietnamese Late Arrival group and the Californian English group was highly significant ($p < .001$ on a Fisher exact probability test), and the difference between the Vietnamese Early Arrival group and the Californian English group was also significant ($p < .02$ on a Fisher exact probability test); however, the difference between the Vietnamese Late Arrival group and the Vietnamese Early Arrival group did not approach significance ($p > .10$ on a Fisher exact probability test.) The distributions of peak pitch classes for the three groups are shown in Figure 4. Given this pattern of results, we conclude

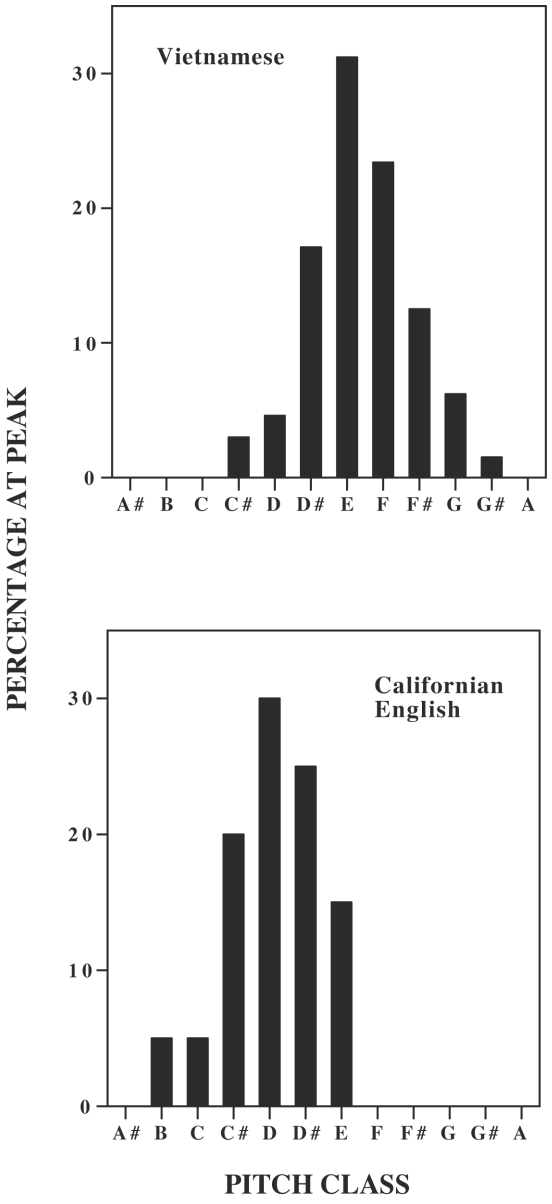

Fig. 3. Distributions of peak pitch classes for two Vietnamese groups combined and for the Californian English group.

that the pitch class template that influences perception of the tritone paradox in adulthood is strongly influenced by the first language to which the individual had been exposed, even if he or she does not speak this language fluently.

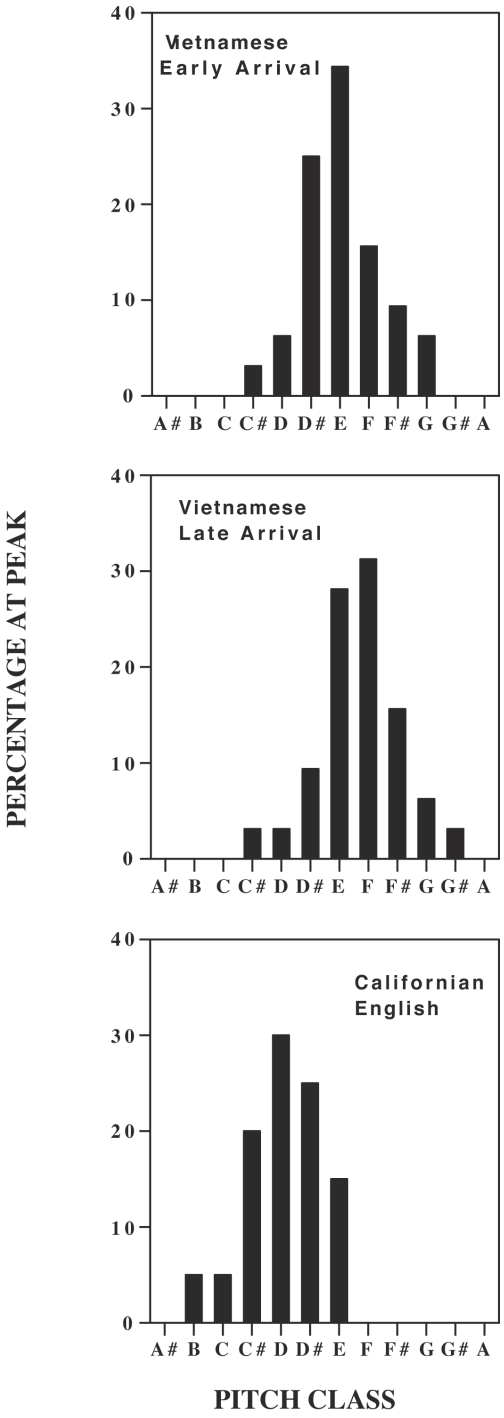

Fig. 4. Distributions of peak pitch classes for the Vietnamese Early Arrival group, for the Vietnamese Late Arrival group, and for the Californian English group.

Experiment 2

In Experiment 1 we found that perception of the tritone paradox varied significantly depending on the language to which the listener had first been exposed. In Experiment 2, we examined the hypothesis that the factor responsible for this association was the pitch range of the speaking voice. Our procedure was similar to that of Deutsch et al. (1990) in a study of monolingual speakers of English. Here we found a significant correlation between a listener's peak pitch classes, as derived from his or her judgments of the tritone paradox, and the octave band containing the largest number of pitch values in a 15-min segment of his or her speech. In the present experiment we inquired whether such a correlation would hold for Vietnamese speakers also.

METHOD

Subjects

Seven subjects took part in this experiment and were paid for their participation. These were two men and five women with a mean age of 46.3 years. All subjects spoke fluent Vietnamese, and none were fluent speakers of English. Six of the subjects had participated in Experiment 1, and one new subject was recruited, who was from North Vietnam.

Procedure and Apparatus

In the first part of the experiment, the subject made judgments of the tritone paradox on two sessions. We averaged the data across the sessions, and so derived the subject's peak pitch classes, using the procedure described earlier. In the case of the six subjects who had participated in Experiment 1, their data from that experiment were employed.

The second part of the experiment consisted of two further sessions. In each session, the subject read out a 5-minute passage from a Vietnamese magazine. The subject's speech was recorded onto DAT tape, and the digitized samples were recorded into a NeXT computer. Pitch estimates were then obtained at 5-ms intervals, using a parallel processing scheme derived from that of Rabiner and Schaffer (1978), with additional signal processing by one of us (M.D.) The pitch estimates were allocated to semitone bins, with the center frequency of each bin determined by the equal tempered scale. Plots were then made of the percentage occurrence of pitch estimates in each bin.

RESULTS AND DISCUSSION

Figure 5 shows, as examples, the distributions of pitch estimates produced by each subject's recorded speech on the first of the two sessions. We note that these plots differ in general shape from those generally produced by speakers of English (see Deutsch et al., 1990, for examples). Vietnamese is a tone language, so the distribution of pitches in the plots shown here reflects the distribution of pitches in Vietnamese tones.

For each subject, we averaged the speech samples across the two sessions, and from this data we derived the octave band containing the largest

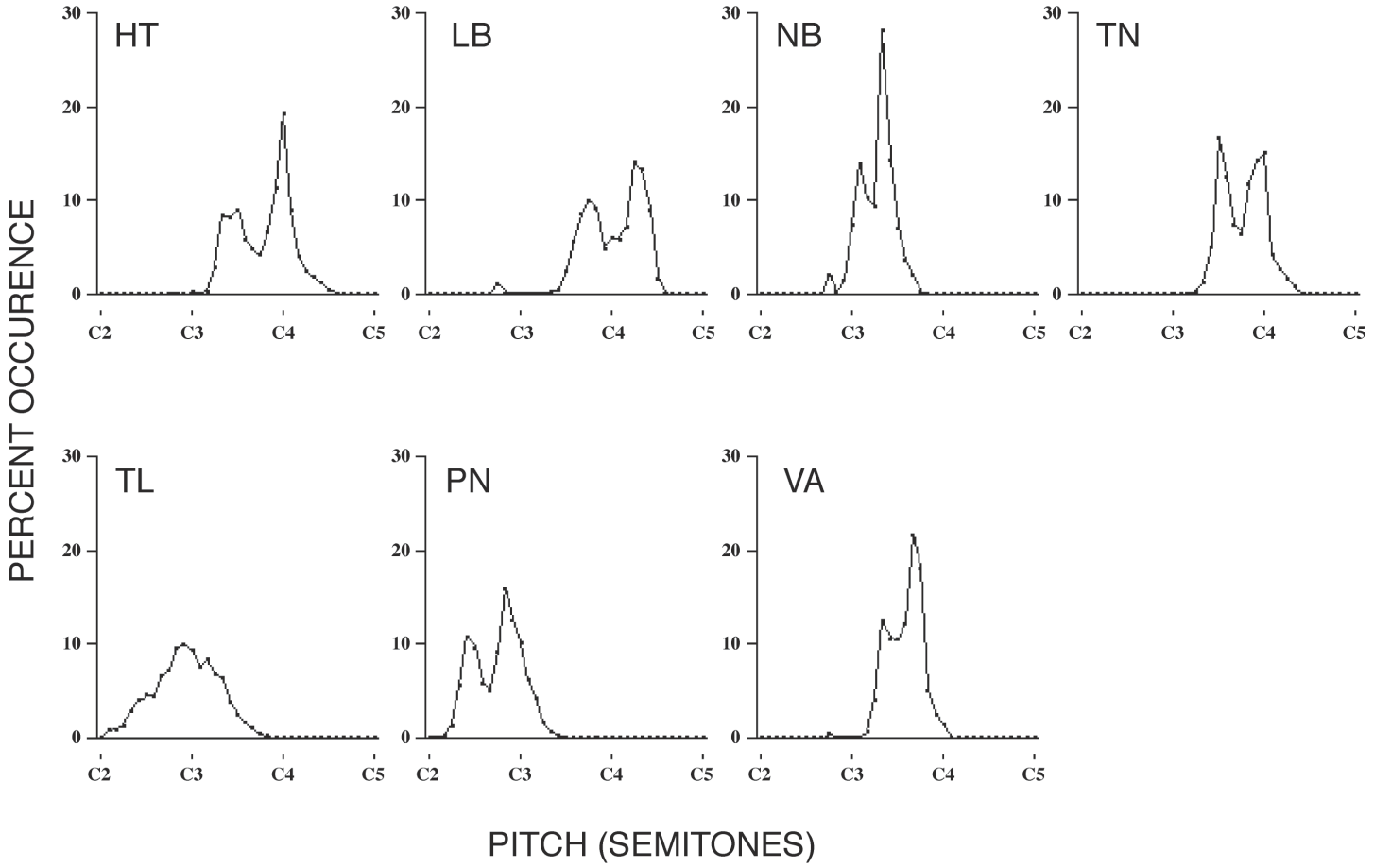

Fig. 5. The percentage occurrence of pitch values in a 5-min segment of each subject's speech, plotted in semitone bins.

number of pitch values in his or her recorded speech. Table 1 shows, for each subject, the pitch classes delimiting this octave band for speech, together with those defining the highest position along the pitch class circle, as determined by his or her judgments of the tritone paradox. It can be seen that, just as in the earlier study by Deutsch et al. (1990), these two values corresponded closely. This correspondence was found to be highly significant ($p < .005$ on a binomial test). The findings therefore strongly suggest that the characteristic of language that influences perception of the tritone paradox is the pitch range of the speaking voice to which the listener has been exposed.

General Discussion

The findings of Experiments 1 and 2 taken together indicate that perception of the tritone paradox can be heavily influenced by a pitch template that is derived from speech heard early in life. Further issues remain to be investigated, however. For example, it is at present unknown whether findings similar to those obtained from the Vietnamese Early Arrival group would be obtained from a group of subjects who were no longer continually exposed to their first language, or from a group who had been born in the United States, and so had received a larger amount of exposure to English in early childhood. It is possible that the influence of speech patterns heard in early childhood might be just as strong in these other groups; however, it might also be attenuated.

More generally, the present study contributes to the literature on critical periods in speech development and the malleability of the perceptual system beyond these critical periods (Johnson & Newport, 1989; Newport, 1990; Newport, Bavelier, & Neville, 2001). Where phonetic aspects of

TABLE 1
Pitch Classes Delimiting Speech Band and Those Defining Highest Position Along the Pitch Class Circle

Subject	Limit of Octave Band for Speech	Highest Position Along Pitch Class Circle	Distance in Semitones
TL	E-F	E-F	0
PN	D♯-E	D♯-E	0
TN	D♯-E	D♯-E	0
LB	F♯-G	F-F♯	-1
VA	C♯-D	D-D♯	+1
NB	G♯-A	G-G♯	-1
HT	D-D♯	F-F♯	+3

speech are concerned, a number of researchers have provided evidence that early bilinguals continue to be influenced by their first language in adulthood (see, e.g., Caramazza, Yeni-Komshian, Zurif, & Carbone ,1973; Pallier, Bosch, and Sebastian-Galles, 1997; Sebastian-Galles and Soto-Faraco, 1999). These researchers have concluded that early bilinguals evidence a lack of plasticity for these elements of speech. However, other researchers have argued instead that the capacity to establish new phonetic categories may remain throughout life (see, e.g., Flege, 1999). The system that is responsible for perception and production of intonation patterns, including pitch range, might parallel the one responsible for the perception and production of phonetic categories; alternatively it might exhibit different developmental characteristics. To our knowledge, this issue has not been directly investigated; however, the present findings strongly indicate that exposure to speech sounds during an early critical period can have a substantial influence on perception of musical patterns that extends into adulthood.[5]

References

Besson, M., & Schon, D. (2003). Comparison between language and music. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 269–293). Oxford: Oxford University Press.

Boe, L., & Rakolofiringa, H. (1975). A statistical analysis of laryngeal frequency: Its relationship to intensity, level, and duration. *Language and Speech*, *8*, 1–13.

Caramazza, A., Yeni-Komshian, G., Zurif, E., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America, 54,* 421–428.

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.

Cutler, A. (1991). Linguistic rhythm and speech segmentation. In J. Sundberg, L. Nord, & R. Carlson (Eds.), *Music, language, speech, and brain* (pp. 157–166). London: Macmillan.

Cutler, A., Dahan, D., & Donselaar, W. van. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*, 141–210.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: Murray.

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.

5. We are grateful to Quyen Doan, Van Doan and Linsey Doan for their help in recruiting subjects and acting as interpreters. The findings described in this article were first presented at the 140th meeting of the Acoustical Society of America (Deutsch, Henthorn, & Dolson, 2000).

Deutsch, D. (1986). A musical paradox. *Music Perception, 3,* 275–280.
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. (1995). *Musical illusions and paradoxes*. La Jolla, CA: Philomel Records.
Deutsch, D. (1996). Mothers and their children hear a musical illusion in strikingly similar ways [Abstract]. *Journal of the Acoustical Society of America, 99,* 2482.
Deutsch, D. (1997). The tritone paradox: A link between music and speech. *Current Directions in Psychological Science, 6,* 174–180.
Deutsch, D., Henthorn, T., & Dolson, M. (2000). Bilingual speakers perceive a musical illusion in accordnce with their first language [Abstract]. *Journal of the Acoustical Society of America*, 108, 2591.
Deutsch, D., Kuyper, W. L., & Fisher, Y. (1987). The tritone paradox: Its presence and form of distribution in a general population. *Music Perception*, *5,* 79–92.
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.
Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barlow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–427). New York, NY: Oxford University Press.
Flege, J. (1999). Age of learning and second-language speech. In D. Birdsong (Ed.) *New perspectives in the critical period hypothesis for second language acquisition* (pp. 101–132). Hillsdale, NJ: Erlbaum.
Giangrande, J. (1998). The tritone paradox: Effects of pitch class and position of the spectral envelope. *Music Perception, 15,* 1–12.
Hanley, T., Snidecor, J., & Ringel, R. (1996). Some acoustic differences among languages. *Phonetica, 14*, 97–107.
Haugen, M., & Chalikia, M. H. (2000). The perception of tritone stimuli by listeners in the Midwest [Abstract]. *Journal of the Acoustical Society of America*, *108*, 2572.
Hollien, H., & Jackson, B. (1973). Normative data on the speaking fundamental frequency characteristics of young adult males. *Journal of Phonetics, 1,* 117–120.
Houtsma, A. J. M. (1995). Pitch perception. In B. C. J. Moore (Ed.), *Hearing* (pp. 267–291). San Diego: Academic Press.
Hudson, A., & Holbrook, A. (1981). A study of the reading fundamental vocal frequency of young black adults. *Journal of Speech and Hearing Research, 24,* 197–201.
Hudson, A., & Holbrook, A. (1982). Fundamental frequency characteristics of young black adults: Spontaneous peaking and oral reading. *Journal of Speech and Hearing Research, 25,* 25–28.
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.
Kunzel, H. (1989). How well does average fundamental frequency correlate with speaker height and weight? *Phonetica, 46,* 117–125.
Majewski, W., Hollien, H., & Zalewski, J. (1972). Speaking fundamental frequency of adult Polish males. *Phonetica, 25,* 19–125.
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.
Pallier, C. Bosch, L., & Sebastian-Galles, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, *64,* B9–B17.
Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*, 717–733.

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

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

Risset, J-C. (1971). Paradoxes de hauteur: Le concept de hauteur n'est pas le même pour tout le monde. *Proceedings of the 7th International Congress on Acoustics, Budapest, S10,* 613–616.

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.

Sebastian-Galles, N., & Soto-Faraco, S. (1999). On-line processing of native and non-native phonemic contrasts in early bilinguals. *Cognition, 72*, 111–123.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36,* 2346–2353.

Spencer, H. (1857). The origin and function of music. *Frazer's Magazine*, *56*, 396–408.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, *55*, 1061–1069.

Terhardt, E. (1991). Music perception and sensory information acquisition: relationships and low-level analogies. *Music Perception*, *8*, 217–239.

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

Yamazawa, H., & Hollien, H. (1992). Speaking fundamental frequency patterns of Japanese women. *Phonetica, 49,* 128–140.