*Music Perception*
Winter 1994, Vol. 12, No. 2, 213–225



# A Regional Difference in Perception of the Tritone Paradox Within the United States

FRANK RAGOZZINE & DIANA DEUTSCH
*University of California, San Diego*

A previous study (Deutsch, 1991) demonstrated a striking difference in perception of the tritone paradox between subjects who had grown up in two different geographical regions. Specifically, a group of subjects who had grown up in California were compared with a group who had grown up in the South of England. When the Californian group tended to hear the pattern as ascending, the English group tended to hear it as descending, and vice versa. This raises the question of whether regional differences also exist within the United States in the way this pattern is perceived. The present study examined the percepts of subjects who had grown up in Mahoning and/or Trumbull counties in Ohio. Two groups were compared: those whose parents had also grown up in this area and those for whom this was not the case. A highly significant difference between these two groups of subjects was obtained, with those in the latter group producing a distribution of percepts similar to that found among Californians and those in the former group producing a different distribution. From this and other analyses of the data, it is concluded that regional differences in perception of the tritone paradox do indeed exist within the United States and that there is in addition an effect of familial background.

THE tritone paradox (Deutsch, 1986, 1987, 1991; Deutsch, Kuyper, & Fisher, 1987; Deutsch, North, & Ray, 1990) is a recently discovered perceptual phenomenon with a number of interesting implications. This paradox is produced by using computer-generated tones consisting only of partials that stand in octave relation; thus pitch class is clearly defined in these tones, but pitch height is ambiguous (Deutsch, 1986; see also Shepard, 1964). Listeners are presented with a sequential pair of such tones that are separated by a half-octave (or tritone), and they judge whether the tones form an ascending or a descending pattern. Surprisingly, listeners perceive this pattern as ascending in some keys, yet as descending in other keys; this violates the principle of equivalence under transposition. More-

Requests for reprints may be sent to either author at the Department of Psychology, University of California, San Diego, La Jolla, CA 92093-0109.

over, when the pattern is presented in any one key, some listeners hear it as ascending, while others hear it as descending instead. Both musicians and nonmusicians experience this paradox, and subjects perceive such patterns in a manner that is consistent over time.

Deutsch has hypothesized a basis for this paradox (Deutsch, 1991; Deutsch et al., 1990). She postulates that one acquires a mental representation of the pitch class circle through a process of developmental learning. The orientation of this mental template is determined by exposure to speech sounds produced by others, with the region of the pitch class circle perceived as highest corresponding to the pitch classes delimiting the octave band for speech that is most prevalent in the listener's linguistic community.[1] The proposal that a template acquired through experience plays a role in pitch perception has also been made by others (Thurlow, 1963; Whitfield, 1967, 1970). In particular, Terhardt (1974, 1991) has proposed that we attribute a fundamental from a series of harmonics as the result of a learned association between different components of the harmonic series and that exposure to speech plays an important role in this learning process.

A pitch class template may provide an evolutionary advantage, in that it enables the emotional state of a speaker to be identified within a given linguistic community regardless of whether the speaker is male or female. When presented with tones that are ambiguous with respect to height, and in the absence of other cues on which to base judgments of relative height, subjects rely on this pitch class template in order to make their judgments. In this sense, everyone possesses a form of absolute pitch. This finding is consistent with the suggestion that some form of absolute pitch is more prevalent in the general population than had traditionally been believed (Terhardt and Seewann, 1983; Terhardt & Ward, 1982).

Two lines of evidence provide support for this model. First, the pitch classes that listeners perceive as highest in this paradigm correlate with those delimiting the octave band that contains the largest number of pitch values in their spontaneous speech (Deutsch et al., 1990). Second, a geographical correlate with perception of the tritone paradox has been found, with subjects from the South of England tending to perceive the pattern in a fashion that is opposite that generally found among Californians. Specifically, on the basis of an earlier study of undergraduates at the University of California, San Diego (UCSD) (Deutsch et al., 1987), Deutsch hypothesized that the orientation of the pitch class circle for Californians would be such that those pitch classes that fell in the half of the circle moving clockwise from A♯-E would be heard as highest (Figure 1, top). As can be

1. The octave band for speech is defined as the octave range containing the largest proportion of fundamental frequencies in the listener's spontaneous speech. See Dolson (1994) for an evaluation of the literature concerning pitch range for speech in relation to linguistic community.

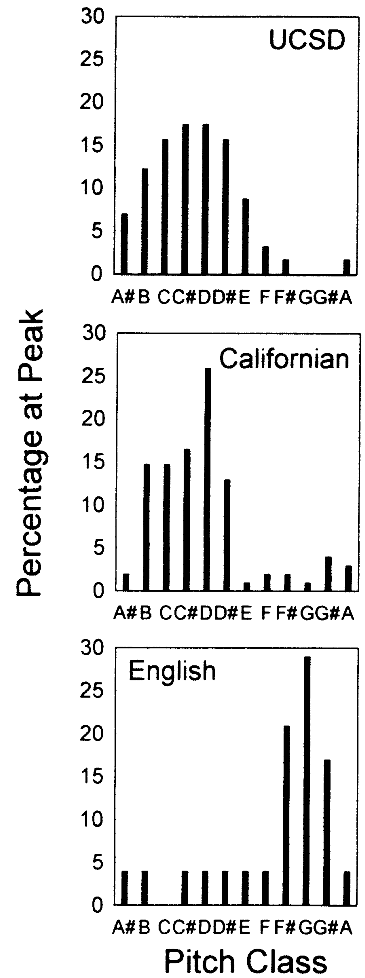


**Fig. 1.** Distributions of peak pitch classes within Californian and Southern English subject populations, found in the study of Deutsch (1991). Also shown for comparison is the distribution of peak pitch classes obtained by Deutsch, Kuyper, and Fisher (1987) within a population of subjects who were undergraduates at UCSD, most of whom are assumed to have grown up in California.

seen in Figure 1 (middle), this hypothesis was confirmed. However, the English subjects produced a different pattern of results. These subjects tended to hear pitch classes from the opposite half of the pitch class circle (i.e., pitch classes moving clockwise from E-A#) as higher (Figure 1, bottom). As will be described later, it appears that the "Californian" mode of perception is the most prevalent one in the United States.

The purpose of the present experiment was to explore the possibility that regional differences in perception of the tritone paradox might exist within the United States. One of the authors (FR) perceives these tritone patterns so that the pitch classes that are heard as highest (termed *peak pitch classes*) are F and F♯. This type of percept occurs frequently among subjects from the South of England, but is rare among Californians (Deutsch, 1991).

F.R. had grown up entirely within Mahoning and Trumbull counties, which are located adjacent to each other in Northeastern Ohio, and both of his parents had also grown up in these counties. It was hypothesized, therefore, that subjects who had grown up in this area, and whose parents had also grown up in this area, might as a group hear this pattern in a fashion that was different from Californians.

More specifically, one may ask what aspects of the linguistic subculture influence the orientation of the pitch class circle for a given listener. Individuals are exposed to multiple sources of linguistic influence from parents, peers, and media such as television and radio. All these factors may influence the orientation of an individual's pitch class template. On the basis of pilot work, we hypothesized that the speech of the listener's parents might play a major role in determining the orientation of the pitch class template. We therefore decided to take into account specific effects of familial background on perception of these patterns. In this way we obtained a "pure" group of subjects who had not only grown up in Mahoning and Trumbull counties, but whose parents had both grown up in this area also. We expected to find that a group of subjects who had grown up in this area, and whose parents had also grown up in this area, would have a distribution of peak pitch classes that was significantly different from that of a group who had grown up in this area, but whose parents had grown up elsewhere within the United States.

## Method

### STIMULUS PATTERNS

The stimulus patterns, and the method for generating them, were identical to those used by Deutsch et al. (1987). Each tone consisted of six octave-related sinusoids, the amplitudes of which were determined by a fixed, bell-shaped spectral envelope. The equation describing the envelope is

$$A(f) = 0.5 - 0.5 \cos\left[\frac{2\pi}{\gamma}\log\beta\left(\frac{f}{f_{min}}\right)\right],$$

where $f_{min} \le f \le \beta\gamma f_{min}$. $A(f)$ is the relative amplitude of a sinusoid with a frequency of $f$ Hz, $\beta$ is the frequency ratio between adjacent sinusoids, $\gamma$ is the number of sinusoids making up the tone, and $f_{min}$ is the minimum nonzero frequency. In the present experiment, six octave-related components were used. Thus, $\beta = 2$, $\gamma = 6$, and the spectral envelope spanned six octaves from $f_{min}$ to $64f_{min}$.

Each tone pair was generated under each of four spectral envelopes, which were spaced at half-octave intervals. These envelopes were centered at $C_4$ (262 Hz, $f_{min}$ = 32.7 Hz), $F\#_4$ (370 Hz, $f_{min}$ = 46.2 Hz), $C_5$ (523 Hz, $f_{min}$ = 65.4 Hz), and $F\#_5$ (740 Hz, $f_{min}$ = 92.4 Hz). Figure 2 shows the relative amplitudes of the components for the tritone pair G -D generated under each of these four envelopes. Note that the amplitudes of the components of pitch class G generated under the $C_4$ and $C_5$ envelopes were identical to the amplitudes of the components of pitch class D generated under the $F\#_4$ and $F\#_5$ envelopes, and vice versa. This relationship held true for all tritone pairs. By averaging over these four envelopes, we counterbalanced for possible spectral effects due to the relative amplitudes of the components and also obtained a profile of the paradox with the positions of the tones along the height continuum varying over a two-octave range.

Twelve tritone pairs were generated under each of the four spectral envelopes. The pitch classes of these pairs were 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. For each of the four envelopes, four quasi-random orderings of the 12 tone pairs were constructed, with the restriction that the same pitch classes did not occur in any two consecutive pairs. Tones were presented in 16 blocks, with each block consisting of all 12 of the tritone pairs generated under one of the four spectral envelopes in one of the four orderings.

## PROCEDURE

Subjects were tested individually and were asked to judge whether each tone pair formed an ascending or descending pattern. All tones had a duration of 500 ms, and there were no gaps between tones within a pair. Within blocks, tone pairs were separated by 4-s intertrial intervals. Blocks were separated by 1-min pauses, and a 5-min break was given between the eighth and ninth blocks. Each subject was tested in the same manner in a

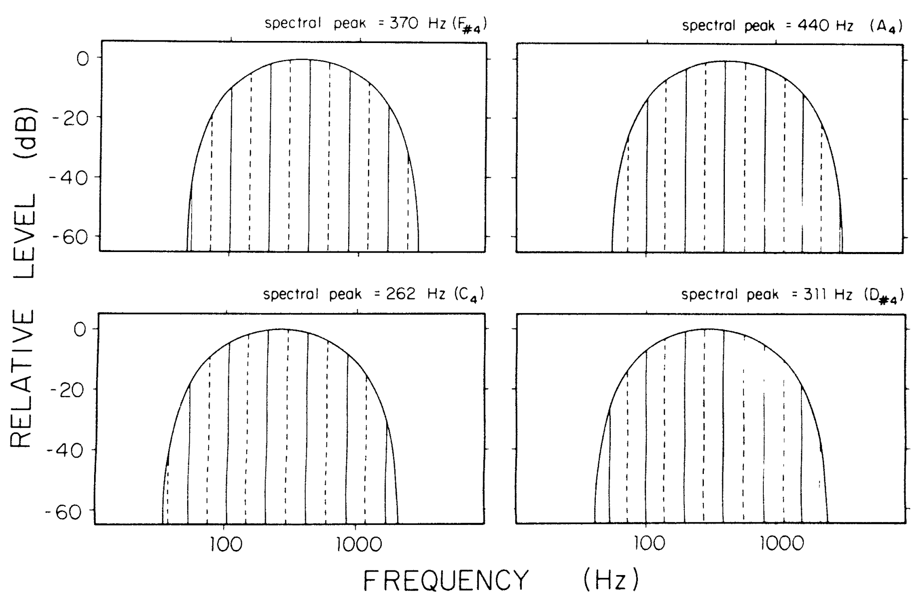


**Fig. 2.** Representation of the spectral composition of the tones comprising the D-G# tritone pattern, generated under the four spectral envelopes used in the study. Dashed lines represent tones of pitch class D, and solid lines represent tones of pitch class G#. The two sets of spectra are superimposed in these diagrams; however, the tones were presented in succession.

second session, which was held on a different day, and responses from both sessions were averaged. All testing took place in the Mahoning and Trumbull county area of Ohio.

Before the testing, subjects filled out a questionnaire detailing where they had lived until age 20, and also where their parents had grown up during their first 20 years of life. This questionnaire also inquired concerning the subjects' musical training.

### EQUIPMENT

Tones were generated on a VAX 11/780 computer using the cmusic sound synthesis software and were recorded onto a Sony digital audio tape using a portable Sony DAT player/recorder, model TCD-D10PRO. Stimuli were presented binaurally through Grason-Stadler headphones at approximately 72 dB SPL.

### SUBJECTS

Twenty-eight subjects who had spent the majority of their first 20 years of life as residents of Mahoning and/or Trumbull counties in Ohio participated in the experiment. Seventeen of these subjects were categorized as "local," based on the fact that both parents had also grown up in Mahoning and/or Trumbull counties. The remaining 11 subjects were categorized as "alien," based on the fact that at least one parent had grown up elsewhere within the United States. The age range of the subjects was 18–33 years, with a mean of 24.25 years.

Subjects were selected on the basis of making no more than four errors out of 48 in a preliminary test in which they judged whether pairs of pure tones that were related by a tritone formed ascending or descending patterns. All potential subjects also underwent audiometry, and those who demonstrated hearing loss did not participate in the experiment. All subjects were paid for their participation.

## Results and Discussion

For each subject, the percentage of judgments that a tritone pair formed a descending pattern, averaged across the four spectral envelopes, was plotted as a function of the first tone of the pair. As can be seen in the examples shown in Figure 3, there was a strong influence of pitch class on these judgments, and the direction of this influence varied from subject to subject.

In order to determine the orientation of the pitch class circle for each subject, and thus his or her peak pitch classes, we used the following procedure: We bisected the pitch class circle so as to maximize the difference between the upper and lower halves, based on the number of times the subject heard each ordered pitch class pair as descending. We then oriented the circle so that the line of bisection was horizontal. This provided an orientation of the pitch class circle for the subject. The two pitch classes at the top of the circle were then taken as the peak pitch classes.[2]

2. This procedure was identical to that used by Deutsch et al. (1987) and Deutsch (1991).

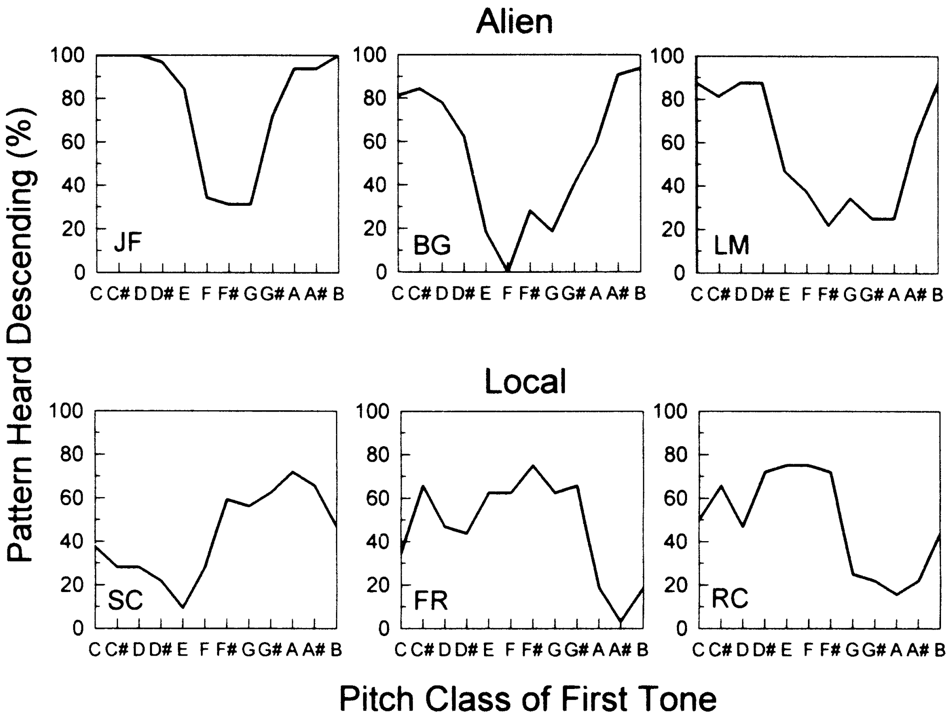


**Fig. 3.** Percentage of judgments that a tone pair formed a descending pattern, plotted as a function of the pitch class of the first tone of the tritone pair, averaged over the four spectral envelopes. Top row: Functions for three "alien" subjects. Bottom row: Functions for three "local" subjects.

Thus, for example, the peak pitch classes of "alien" subject BG, whose data are shown in Figure 3, were C and C#, and the peak pitch classes of "local" subject SC, whose data are also shown in Figure 3, were G# and A. The orientations of the pitch class circle with respect to height for these two subjects are shown in Figure 4.

The distributions of peak pitch classes were then plotted for the "alien" and "local" groups separately (Figure 5). Consistent with earlier findings, the peaks of the "alien" group were entirely in the half of the pitch class circle moving clockwise from A# to E. However, the peaks of the "local" group showed a different distribution, with nine subjects having peaks in the same half of the circle as did the "aliens" and eight of these subjects having peaks in the opposite half of the circle. This difference between the two groups in the distribution of peaks was statistically significant ($p < .008$ on a Fisher exact probability test).

Separate analyses were conducted to search for possible effects of age, sex, or musical training. There was no significant difference in the distribution of peak pitch classes for the 14 older versus the 14 younger subjects,

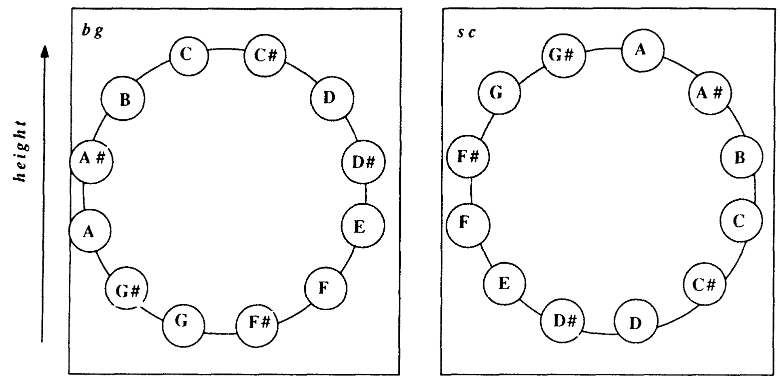


**Fig. 4.** Orientations of the pitch class circle with respect to height, derived from the judgments of "alien" subject BG and "local" subject SC, whose data are plotted in Figure 3. For subject BG, the peak pitch classes were C and C♯, and for subject SC, they were G♯ and A.

neither were there significant differences in the distribution of peaks for the older versus the younger subjects within the "alien" and "local" groups analyzed separately (in each case, $p > .05$ on a Fisher exact probability test). An analysis of the distribution of peaks for male versus female subjects also produced no significant difference, both for the entire group of subjects, and also for the "alien" and "local" groups taken separately (in each case, $p > .05$ on a Fisher exact probability test.) Finally, subjects were designated as musically trained if they had received more than 2 years of formal training. A comparison of "trained" with "untrained" subjects also failed to reveal a significant difference either for the group as a whole, or for the "alien" and "local" groups taken separately (in each case, $p > .05$ on a Fisher exact probability test).

The present findings show that a regional difference in perception of the tritone paradox within the United States does indeed exist. Subjects categorized as "local" perceived these tritone patterns in a manner that was significantly different from subjects categorized as "alien." Among the "alien" group, all peak pitch classes were confined to B, C, C♯, D, and D♯ (i.e., they were in the "Californian" range). However, the peak pitch classes of the "local" group formed an entirely different distribution.

Another interesting finding involves a further analysis of the distribution of peak pitch classes among the "local" group. Roughly half of these subjects had peak pitch classes in the "Californian" range (and so in the range that was also typical of the "alien" group), however, the other half had peak pitch classes in the opposite range. In addition, the profiles from the "local" subjects were, in general, less pronounced than were those from the "alien" subjects. One may therefore hypothesize that factors other than their parents' speech had influenced the pitch class

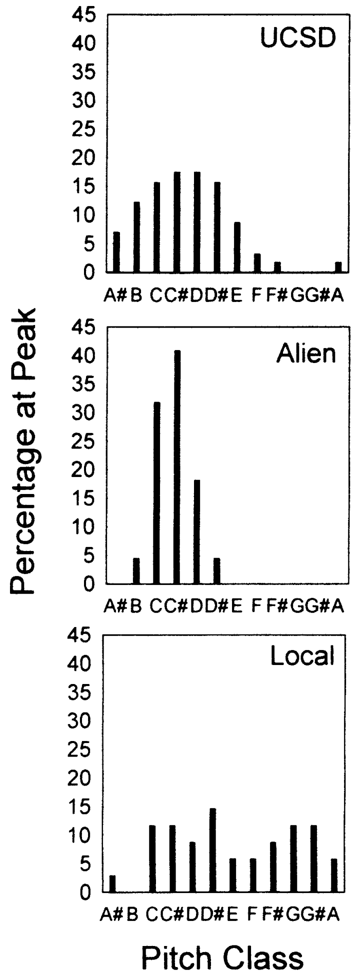


**Fig. 5.** Distributions of peak pitch classes within the "alien" and "local" subject populations. Also shown for comparison is the distribution of peak pitch classes obtained by Deutsch, Kuyper, and Fisher (1987) within a population of subjects who were undergraduates at UCSD.

templates of these subjects. More specifically, one may speculate that there is a canonical manner of perceiving these sound patterns within the United States, which is propagated by the linguistic influences of media such as radio and television, with the result that the peak pitch classes tend to be those found in the earlier studies on Californians. On the basis of previous formal and informal studies (involving hundreds of subjects altogether), we have found that the most probable percept among American subjects is such that the peak pitch classes are in the "Californian" range. For example, this mode of perception was also found in a study by Deutsch et al. (1987) of UCSD undergraduates, and in a study by

Deutsch and Miyoshi (in preparation) on subjects who had grown up in the greater Boston area. However, it seems that various geographical regions exist where people still retain an orientation of the pitch class circle that differs from the canonical one. If the above line of reasoning is correct, then our "local" subjects whose peaks lie in the "Californian" range may for some reason be more susceptible to outside linguistic influences in establishing an orientation of the pitch class template, while those whose peaks lie in the opposite range may be less susceptible to these factors. From another perspective, this bimodal distribution of peak pitch classes among the "local" subjects may reflect an ongoing process in which this particular geographical area is gradually switching over to the canonical percept.

If the foregoing line of reasoning is correct, then one may hypothesize that the "local" subjects, as a group, were invoking two templates; a canonical template with a peak in the "Californian" range, and a "local" template with a peak roughly in the opposite half of the pitch class circle. Additional evidence for this hypothesis comes from an examination of the distribution of peak pitch classes within the "alien" and "local" groups, determined separately for tones generated under different spectral envelopes. We compared the distributions for tones generated under envelopes centered on $C_4$ and $F\sharp_4$ (i.e., the two lower envelopes) with those generated under envelopes centered on $C_5$ and $F\sharp_5$ (i.e., the two higher envelopes). As shown in Figure 6, for the alien group, the distributions of peak pitch classes generated under these two categories of envelope were quite similar (although they were more strongly clustered under the lower envelopes). However, for the local group, these distributions were instead quite different, with peak pitch classes C♯, D, and D♯ occurring most often under the lower envelopes and peak pitch classes G and G♯ occurring most often under the higher envelopes instead.

This pattern of results is consistent with the hypothesis that the "local" subjects, as a group, were basing their judgments rather strongly on two alternative templates. The overall heights of the tones influenced which of these two templates was used in making any particular judgment. When presented with tones generated under the lower envelopes, there was an increased tendency for these subjects, as a group, to make judgments based on the canonical American template. However, when presented with tones generated under the higher envelopes, there was an increased tendency for these subjects to make judgments with reference to the "local" template, which has the opposite orientation.

We may speculate that this pattern of results reflects differences in the pitch ranges of the speaking voices of the two groups of subjects. Some "local" subjects may have an octave band for speech that is higher in pitch than that of the "alien" subjects, so that they compare tones generated

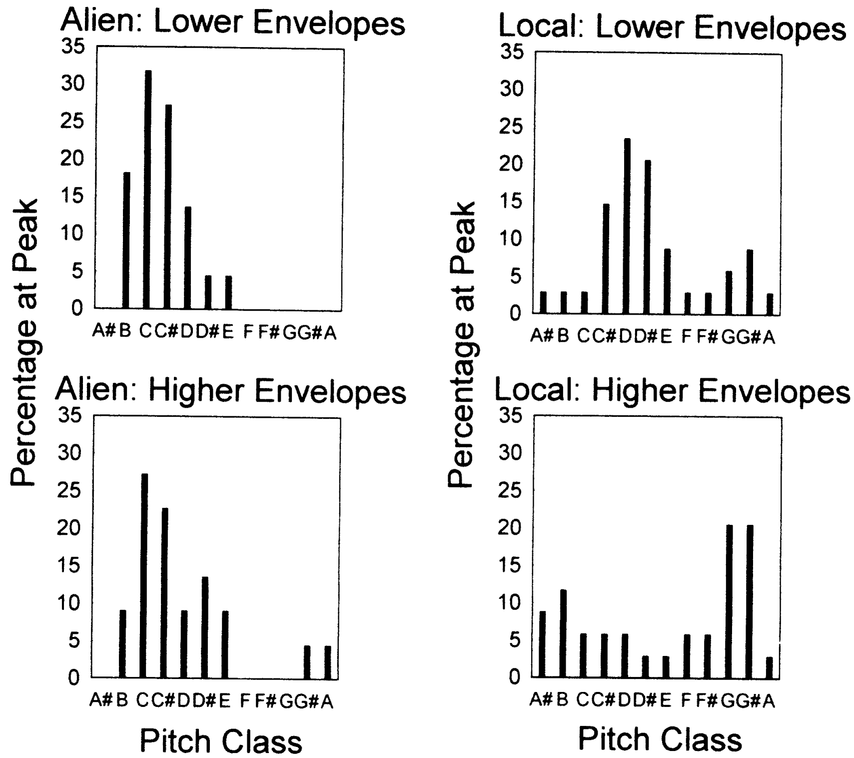


**Fig. 6.** Distributions of peak pitch classes within the "alien" and "local" populations, plotted separately for tones generated under envelopes centered on $C_4$ and $F\sharp_4$ ("lower envelopes") and for tones generated under envelopes centered on $C_5$ and $F\sharp_5$ ("higher envelopes").

under the higher envelopes against their more appropriate "local" template. However, they may also compare tones generated under the lower envelopes against the canonical American template (see also Deutsch, 1994, for a discussion of this issue).

A further analysis supports this hypothesis. For each subject, it was determined whether the pitch class circle could be cleanly divided such that none of the scores in the upper half of the circle were lower than any of the scores in the lower half. This analysis was performed separately for tones generated under the lower and higher spectral envelopes. We then examined the judgments of those subjects who had met this criterion for tones generated under only one of these two sets of envelopes (either the lower or the higher one), but who failed to meet the criterion for tones generated under the other set of envelopes. Of these subjects, those whose peak pitch classes were in the "Californian" range tended to meet the criterion for tones generated under the lower envelopes, whereas those who peak pitch classes were in the opposite range tended to meet the criterion for tones generated under the higher envelopes instead. Ten of the 13 subjects who fell into this category showed this pattern. These findings are consistent with those of Deutsch (1994) in a further analysis of the data obtained by Deutsch (1991) from Californian and English subjects.[3]

In sum, the present study shows that regional differences in perception of the tritone paradox do exist within the United States. It remains to be seen which other areas of the country exhibit similar, or even more extreme, differences.[4–6]

## References

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

Deutsch, D. The tritone paradox. Effects of special variables. *Perception & Psychophysics,* 1987, *41,* 563–575.

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

Deutsch, D. The tritone paradox: Some further geographical correlates. *Music Perception, 1994, 12,* 125–136.

Deutsch, D., Kuyper, W. L., & Fisher, Y. The tritone paradox: Its presence and form of distribution in a general population. *Music Perception,* 1987, *5,* 79–92.

Deutsch, D., & Miyoshi, D. Perception of the tritone paradox among subjects in the Boston area. (in preparation)

Deutsch, D., North, T., & Ray, L. The tritone paradox: Correlate with the listener's vocal range for speech. *Music Perception,* 1990, *7,* 371–384.

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

Ragozzine, F., & Deutsch, D. A regional difference within the United States in perception of the tritone paradox. *The Journal of the Acoustical Society of America,* 1993, *94* (3, Pt. 2), 1860. (abstract)

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

3. Further analyses, analogous to those described in Deutsch (1994) were also performed. The data for subjects whose overall peak pitch classes were in the "Californian" range were compared with data for those whose overall peak pitch classes were in the opposite region of the pitch class circle (regardless of familial background). Consistent with Deutsch's findings, subjects in the former category tended to produce more pronounced profiles relating pitch class to perceived height for tones generated under the lower envelopes (i.e., that peaked at $C_4$ and $F\sharp_4$) than under the higher ones (i.e., that peaked at $C_5$ and $F\sharp_5$). In contrast, subjects in the latter category showed an overall tendency to produce more pronounced profiles for tones generated under the higher envelopes than under the lower ones. Also consistent with the analysis of Deutsch (1994), for subjects in the former category, the peak pitch classes were more strongly clustered for tones generated under the lower envelopes than under the higher ones. In contrast, for subjects in the latter category, the peak pitch classes were more strongly clustered for tones generated under the higher envelopes than under the lower ones.

4. This work was first reported in abstract form by Ragozzine and Deutsch (1993).

5. The authors thank the members of the psychology department of Youngstown State University, and in particular Dr. Joyce Segreto, for the cooperation and assistance provided in support of this research.

6. Preparation of this paper was supported in part by National Institute of Mental Health Training Grant 5T32MH14268.

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

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

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

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

Thurlow, W. R. Perception of low auditory pitch: A multicue mediation theory. *Psychological Review,* 1963, *70,* 515–519.

Whitfield, I. C. *The auditory pathway.* London: Arnold, 1967.

Whitfield, I. C. Central nervous processing in relation to spatio-temporal discrimination of auditory patterns. In R. Plomp & G. F. Smoorenberg (Eds.), *Frequency analysis and periodicity detection in hearing.* Leiden: Sijthoff, 1970, pp. 136–146.