Music Perception
Fall 1994, Vol. 12, No. 1, 125–136

Notes & Comment

The Tritone Paradox: Some Further Geographical Correlates

DIANA DEUTSCH
University of California, San Diego

In a study by Deutsch (1991), a large and highly significant difference in perception of the tritone paradox was found between a group of subjects who had grown up in California and a group who had grown up in the south of England: In general, where the Californian group tended to hear the pattern as ascending the English group tended to hear it as descending, and vice versa. The present paper documents some further geographical correlates that are derived from the data obtained by Deutsch (1991). The strength of the relationship of pitch class to perceived height was found to depend on the overall heights of the spectral envelopes under which the tones were generated. However, the direction of this dependence differed significantly depending on the subject population. For subjects showing a "Californian pattern" (i.e., whose overall peak pitch classes were in the range moving clockwise from A♯–B to D♯–E), this relationship was more pronounced for tones generated under lower spectral envelopes, and so when the tones were perceived as lower in overall height. In contrast, for subjects showing an "English pattern" (i.e., whose overall peak pitch classes were in the opposite region of the pitch-class circle), this relationship was more pronounced for tones generated under higher spectral envelopes, and so when the tones were perceived as higher overall instead. Given the literature on the pitch of speech as a function of linguistic community, these findings provide further evidence that perception of the tritone paradox is related to the processing of speech sounds.

THE tritone paradox, which was first described by Deutsch (1986), is generated by two successively presented tones that are related by a half-octave. Each tone consists of six octave-related harmonics whose amplitudes are determined by a bell-shaped spectral envelope. Such tones are clearly defined in terms of pitch class—for example, one tone pair is

Requests for reprints may be sent to Diana Deutsch, Department of Psychology (0109), University of California, San Diego, La Jolla, CA 92093.

clearly heard as C followed by F♯, and another as G♯ followed by D. However, the tones are ambiguously defined in terms of height—for example, a tone pair that consists of C followed by F♯ can in principle be heard either as ascending or as descending.

When listeners are asked to determine whether such tone pairs form ascending or descending patterns, their judgments are systematically related to the positions of the tones along the pitch-class circle: Tones in one region of the circle are heard as higher and those in the opposite region as lower. Even more remarkably, there are strong differences among listeners as to which tone pairs are heard as ascending and which as descending, and so as to which region of the pitch-class circle is heard as higher and which as lower. Thus, for example, some listeners hear the pattern C–F♯ as ascending and F♯–C as descending, whereas other listeners obtain the opposite percepts. As illustration, Figure 1 shows one subject's judgments, together with the orientation of the pitch-class circle with respect to height derived from these. The pitch classes that are perceived as highest along this circle are termed *peak pitch classes*.

In considering possible explanations of these phenomena (see also Deutsch, 1987; Deutsch, Kuyper, & Fisher, 1987), the author conjectured that perception of the tritone paradox might be related to the processing of speech sounds. More specifically, it was hypothesized that during childhood, listeners acquire a long-term representation of the pitch range of the speaking voices to which they are most frequently exposed. This representation includes a definition of the octave band that contains the largest

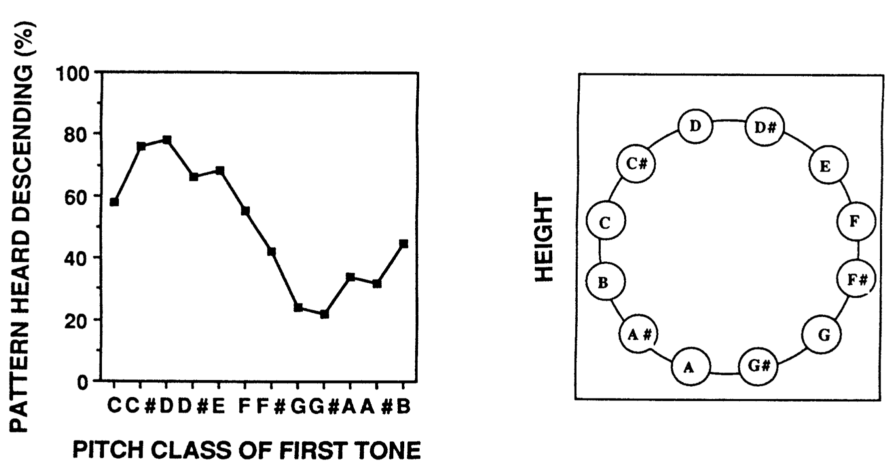

Fig. 1. Graph on left shows, for one subject, the percentages of judgments that a tone pair formed a descending pattern, plotted as a function of the pitch class of the first tone of the pair. Illustration on right shows the orientation of the pitch-class circle with respect to height derived from these judgments. For this subject, the peak pitch classes were D and D♯.

proportion of pitch values in such speech, and this in turn influences the pitch range of the listener's own speaking voice (Deutsch, 1991; Deutsch, North, & Ray, 1990). Furthermore, the pitch classes delimiting the octave band are taken by the listener as defining the highest position along the pitch-class circle, which in turn defines the orientation of the circle with respect to height as reflected in judgments of the tritone paradox.

Support for the conjecture that perception of the tritone paradox is related to the processing of speech sounds was found by Deutsch et al. (1990), who obtained samples of roughly 15 min of spontaneous speech from a group of subjects. The pitch classes delimiting the octave band containing the largest number of pitch values in these speech samples were found to correlate significantly with the pitch classes that stood at the highest position along the circle, as derived from judgments of the tritone paradox.

Support for the conjecture that such a pitch-class template is acquired developmentally through exposure to speech sounds in the listener's environment was obtained by Deutsch (1991) in a study comparing perception of the tritone paradox among a group of Californians and among a group from the south of England: When the Californians tended to hear the pattern as ascending, the English tended to hear it as descending, and vice versa. Further supporting evidence was provided by Dolson (1994) in a review of the literature on the pitch of speech as a function of linguistic community. As he pointed out, most listeners confine the pitch range of their speech to roughly an octave, and the pitch of women's voices in a given linguistic community is close to an octave above that of men. Within a given community, most speakers have pitch ranges that are within three semitones of the group mean (or even narrower). However, when comparison is made between communities, large differences between group means can be demonstrated. For example, a difference of nearly five semitones has been found between the group means for Californian and for Polish males. Also very significantly, the pitch range of speech is surprisingly independent of the physiological characteristics of the speaker, such as laryngeal size or various body dimensions such as height, weight, and chest size.

The above hypothesis raises an interesting question. Assuming that there is an agreed-upon template that is used in interpreting speech within a given community, how does the listener deal with the pitch characteristics of speech from different communities? This question leads us to hypothesize further that listeners develop not only a primary pitch-class template that is appropriate to the dialect that they most frequently encounter, but also one or more subsidiary templates that are appropriate to dialects that they encounter less frequently (or possibly have encountered at a less critical age).

Applying this line of reasoning to the tritone paradox, we can hypothesize that for subjects within a given linguistic community, tones in one pitch range might produce a pronounced profile relating pitch class to perceived height, while tones in different pitch ranges might produce other, subsidiary profiles.

The profiles in the experiment of Deutsch (1991) were obtained by using tone pairs that were created under envelopes placed at four different positions along the spectrum. The envelopes were spaced at half-octave intervals so as to control for possible effects based on the relative amplitudes of the individual harmonics of the tones. As shown in Figure 2, the peaks of the spectral envelopes were 262 Hz (C_4), 370 Hz ($F\sharp_4$), 523 Hz (C_5) and 740 Hz ($F\sharp_5$).

In this experiment, the data from tones generated under the four enve-

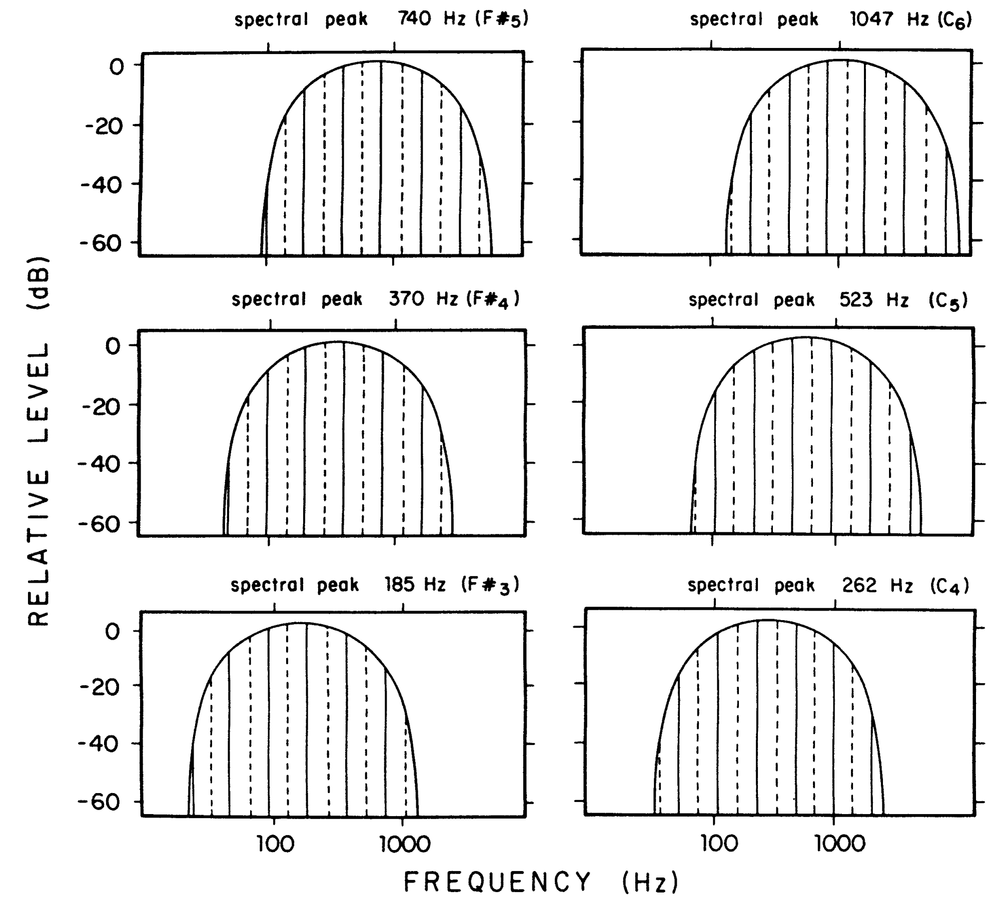

Fig. 2. Represention of the spectral composition of the tones constituting the D–G♯ pattern, when generated under the four envelopes used in the study of Deutsch (1991). Dashed lines indicate tones of pitch class D and solid lines indicate tones of pitch class G♯. In the experiment, the tones D and G♯ were presented in succession, although the spectra for these two tones are superimposed in the illustration.

lopes were averaged so as to derive the overall peak pitch classes for each subject. As shown in Figure 3, a strong effect of geographical region in this averaged data was found: The overall pitch class template was shown to be oriented by Californians and by southern English in quite different ways.

Given the considerations advanced above, the present paper explores the possibility that the subjects in the Deutsch (1991) study may have been invoking more than one pitch-class template, depending on the overall heights of the envelopes under which the tones were generated, and so

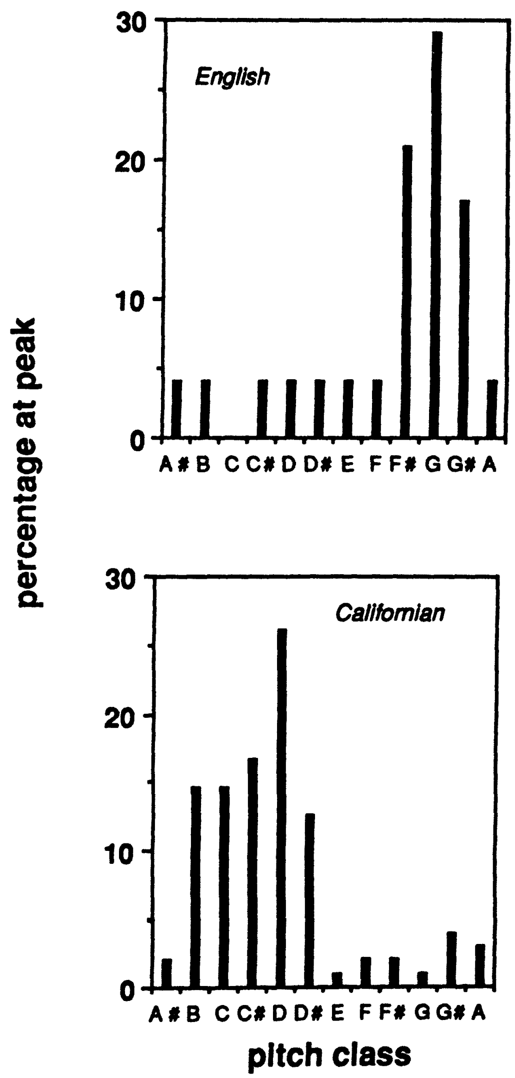

Fig. 3. Distributions of peak pitch classes within the English and Californian subject populations in the experiment of Deutsch (1991), with data averaged over tones generated under all four spectral envelopes. Reprinted from Deutsch (1991).

depending on the perceived overall heights of the tones. More specifically, it appears from the existing literature that the pitch range for speech among Californians is quite low (see, for example, Hanley, Snidecor, & Ringel, 1966). We may therefore conjecture that the Californian subjects would have tended to produce strong profiles relating pitch class to perceived height for tones generated under the lower envelopes used in this study, together with weaker, subsidiary profiles for tones generated under the higher envelopes.

In contrast, British English appears to be characterized by rather wide pitch excursions, particularly at the high end (see, for example, Collier, 1991; Willems, Collier, & 't Hart, 1988). Although direct comparison between the two populations has not yet been made, it is reasonable to assume from the existing data that the overall pitch range for British English speech is higher than that for Californian speech. We may then hypothesize that the English subjects would have tended to produce strong profiles for tones generated under the higher envelopes, and in addition weaker, subsidiary profiles for tones generated under lower ones.

To test these predictions, the subjects in the study of Deutsch (1991) were divided into two groups. The first, which we shall describe as showing the "Californian pattern," consisted of the 24 subjects (21 Californians and three English) whose overall peak pitch classes stood in the range moving clockwise from A♯–B to D♯–E. The second, which we shall describe as showing the "English pattern," consisted of 12 subjects (nine English and three Californian) whose overall peak pitch classes were in the opposite region of the circle (see Figure 3).

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 tone of the pair. The data were averaged separately over the two lower envelopes (i.e., those centered at C_4 and $F\sharp_4$) and the two higher ones (i.e., those centered at C_5 and $F\sharp_5$).

Figure 4 displays the results of this analysis, plotted separately for subjects showing the "Californian pattern" and for those showing the "English pattern." It can be seen that the hypotheses were indeed confirmed. For subjects showing the "Californian pattern," tones generated under the lower envelopes produced a distinct profile, with pitch classes B, C, C♯, D, D♯, and E being heard as higher than the other pitch classes. However, tones generated under the higher envelopes produced no convincing profile relating pitch class to perceived height.

Subjects showing the "English pattern" produced the opposite result. Tones generated under the higher envelopes gave rise to a pronounced profile, with pitch classes F, F♯, G, G♯, A, and A♯ being heard as higher than the others. However, tones generated under the lower envelopes produced a much weaker profile.

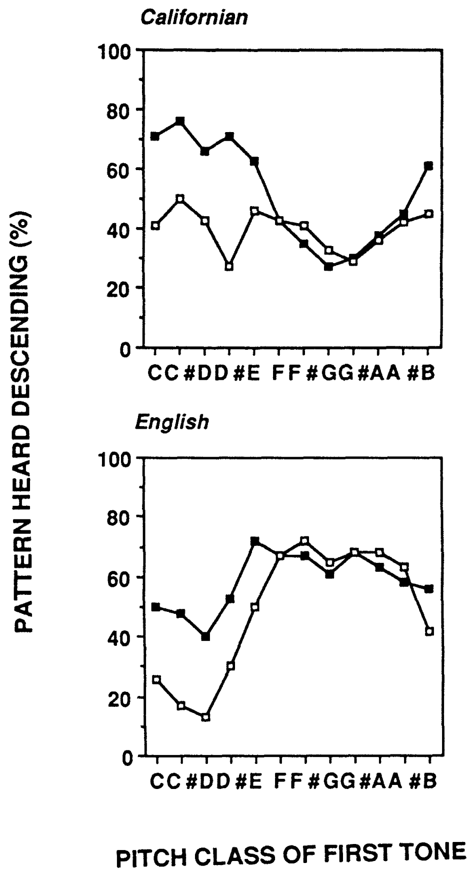

Fig. 4. Percentages of judgments that a tone pair formed a descending pattern, plotted as a function of the pitch class of the first tone of the pair. Filled squares show judgments averaged for tones generated under the lower spectral envelopes, and open squares for tones generated under the higher envelopes. Upper graph shows averaged judgments for subjects who showed an overall "Californian pattern," and lower graph for subjects who showed an overall "English pattern." See text for details.

The possibility arises that, for subjects showing the "Californian pattern," the profiles generated under the higher envelopes were more broadly distributed than were those generated under the lower envelopes. Analogously, for subjects showing the "English pattern," the profiles generated under the lower envelopes might have been more broadly distributed than were those generated under the higher envelopes. The differences in the profiles derived from the pooled data might, therefore, have been due to an artifact of the averaging process.

A further analysis was therefore performed to control for this possibility. The orientation of the pitch-class circle was normalized for each subject, for tones generated under the lower and the higher envelopes sepa-

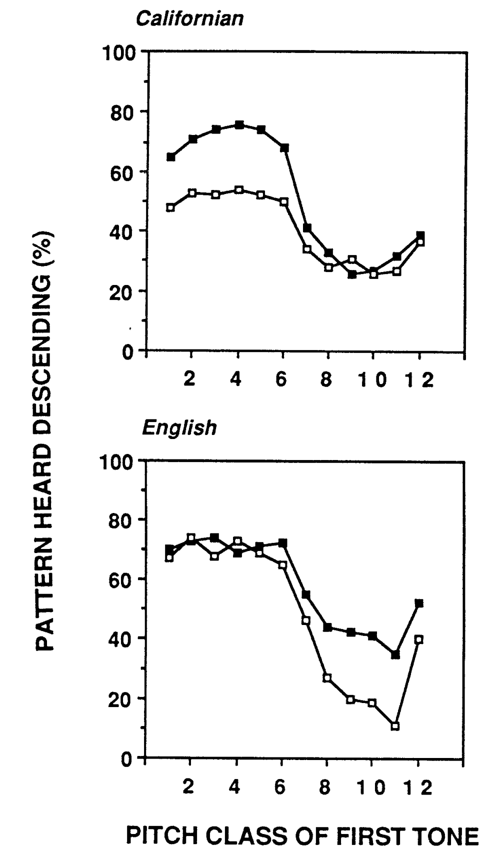

Fig. 5. Percentages of judgments that a tone pair formed a descending pattern, plotted as a function of the pitch class of the first tone of the pair, with the orientation of the pitch-class circle normalized across subjects. Filled squares show averaged judgments for tones generated under the lower spectral envelopes, and open squares for tones generated under the higher envelopes. Upper graph shows averaged judgments of subjects who showed an overall "Californian pattern," and lower graph for subjects who showed an overall "English pattern." See text for details.

rately.[1] The normalization procedure was as follows. First, the pitch-class circle was bisected so as to maximize the difference between the averaged scores in the upper and lower halves. 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 of the circle taking position 1, its clockwise neighbor taking position 2, and so on.

The normalized data, averaged across each group of subjects, are shown in Figure 5. It can be seen that a clear effect of category of envelope was still obtained. For subjects showing the "Californian pattern," the averaged profile was considerably more pronounced for tones generated under

1. This procedure follows that of Deutsch et al. (1987).

the lower envelopes than the higher ones. For the subjects showing the "English pattern," the opposite result was obtained: the averaged profile was considerably stronger for tones generated under the higher envelopes than the lower ones.

In order to make a statistical comparison between the strengths of the profiles under the different categories of envelope for the two groups of subjects, the averaged difference between the scores in the upper and lower halves of the normalized circle was determined for each subject separately. It was found that for 17 of the 24 subjects showing the "Californian pattern," the difference between these scores was greater under the lower envelopes than the higher ones. However, this was true for only two of the 12 subjects showing the "English pattern." The difference between the two groups of subjects on this measure was therefore highly significant ($p < .005$, on a Fisher exact probability test).

A further issue concerns the distributions of peak pitch classes under the two categories of envelope for the two groups of subjects. One might hypothesize that for tones generated under envelopes that produced the more salient profile (i.e., the lower envelopes for subjects showing the "Californian pattern" and the higher envelopes for those showing the "English pattern") the peak pitch classes would be more strongly clustered than for tones generated under the envelopes producing the weaker profile. To examine this issue, the distributions of peak pitch classes were tabulated for each group of subjects separately under each category of envelope.

The results for the subjects showing the "Californian pattern" are displayed in Figure 6. It can be seen that for the lower envelopes, a remarkably orderly and clustered distribution of peak pitch classes was obtained. However, for the higher envelopes, the peak pitch classes were distributed more broadly. The results for the subjects showing the "English pattern" are shown in Figure 7. It can be seen that, in contrast, the peak pitch classes clustered heavily under the higher envelopes but not the lower ones.

To provide another measure of the consistency of the pitch class template for the two groups of subjects under the two categories of envelope, a further analysis was performed. First, it was ascertained for the scores of each subject separately, and for tones generated under each category of envelope, whether the pitch class circle could be bisected such that none of the scores in the upper half of the circle were lower than any of those in the lower half.[2] Of those subjects showing the "Californian pattern," eight showed a consistent profile for the lower envelopes but not the higher ones, and four showed a consistent profile for the higher envelopes but not the lower ones. In contrast, of subjects showing the "English pattern," only one showed a consistent profile for the lower envelopes but not the higher ones, whereas

2. This procedure follows that of Deutsch et al. (1987).

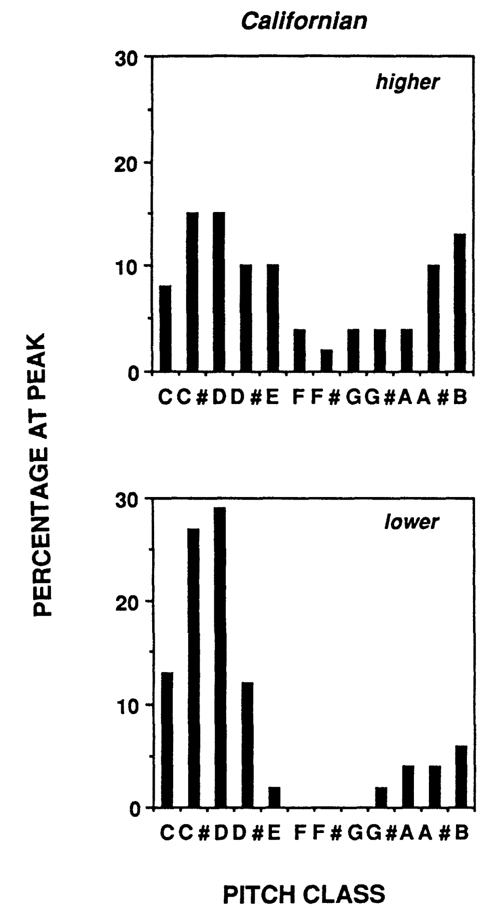

Fig. 6. Distributions of peak pitch classes among subjects who showed an overall "Californian pattern." Upper graph shows peak pitch classes for tones generated under the higher spectral envelopes, and lower graph for tones generated under the lower envelopes. See text for details.

five showed a consistent profile for the higher envelopes but not the lower ones. These trends are in accordance with the results of the other analyses.

Finally, we may inquire into the general size of difference between peak pitch classes for tones generated under the two categories of envelope. It was found that 64% of the subjects in this study showed a difference that was no greater than two semitones, and that the unsigned difference, averaged over all the subjects, was only 1.9 semitones. It is interesting to compare this result with that obtained by Deutsch (1987) using four subjects who had been selected for showing clear profiles relating pitch class to perceived height in judgments of the tritone paradox. As shown by Figures 4 and 5 (p. 567) of that article, one of the subjects showed no difference in peak pitch class between tones generated under envelopes

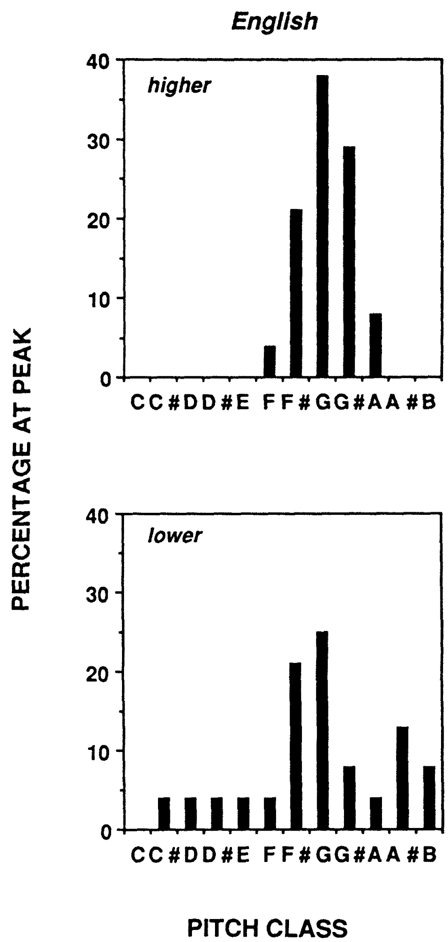

Fig. 7. Distributions of peak pitch classes among subjects who showed an overall "English pattern." Upper graph shows peak pitch classes for tones generated under the higher spectral envelopes, and lower graph for tones generated under the lower envelopes. See text for details.

centered on C_5, $D\#_5$, $F\#_5$, and A_5 on the one hand and tones generated under envelopes centered on C_4, $D\#_4$, $F\#_4$, and A_4 on the other. Two subjects showed a difference of one semitone, and one subject showed a difference of two semitones. It is noteworthy that the present analysis was performed on the data of subjects who had not been selected in this fashion, and that the majority of these subjects nevertheless produced patterns of results that were in the same range.

In summary, this paper has documented some further geographical correlates with perception of the tritone paradox, based on data originally obtained by Deutsch (1991). By using several measures, the salience of the relationship between pitch class and perceived height was found to vary depending on the overall heights of the envelopes under which the tones were generated. For subjects showing the "Californian pattern," this relationship was more salient for tones generated under the lower envelopes used in the study; however, for subjects showing the "English pattern," the relationship was more salient for tones generated under the higher envelopes. Given the existing literature on the pitch of speech as a function of linguistic community, these patterns of relationship provide further evidence that perception of the tritone paradox is related to the processing of speech sounds.

References

Collier, R. Multi-language intonation synthesis. *Journal of Phonetics,* 1991, ***19,*** 61–73.
Deutsch, D. A musical paradox. *Music Perception,* 1986, ***3,*** 275–280.
Deutsch, D. The tritone paradox: Effects of spectral 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., 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., 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.
Hanley, T., Snidecor, J., & Ringel, R. Some acoustic differences among languages. *Phonetica,* 1966, ***14,*** 97–107.
Willems, N., Collier, R., & 't Hart, J. A synthesis scheme for British English intonation. *Journal of the Acoustical Society of America,* 1988, ***84,*** 1250–1261.