Music Perception
Winter 1994, Vol. 12, No. 2, 257–263

Notes & Comment

The Tritone Paradox and the Pitch Range of the Speaking Voice: Reply to Repp

DIANA DEUTSCH
University of California, San Diego

It is shown on a number of grounds that the findings described by Repp (1994) are uninterpretable in relation to those of Deutsch (1991) and Deutsch, North, and Ray (1990). First, the geographical correlates with perception of the tritone paradox obtained by Repp were based on data from subject populations that differed substantially from those studied by Deutsch (1991), and Repp's subject populations were also treated differently from each other. Second, Repp obtained substantial differences in perception of the tritone paradox depending on the spectral envelope under which the tones were generated; however, he attempted to correlate the data obtained from only one of these envelopes with pitch ranges for speech. Third, the procedures used by Repp to determine the pitch ranges for speech were problematic and inappropriate in the context of Deutsch's hypothesis.

In his article, "The tritone paradox and the pitch range of the speaking voice: A dubious connection" (this issue), Repp sets forth an extensive set of comments on the articles by Deutsch, North, and Ray (1990) and Deutsch (1991). The present article replies to these comments. It is shown that the relationships between Repp's findings and those of Deutsch and coworkers are uninterpretable on two general grounds. First, Repp used procedures that differed in several critical respects from those of Deutsch and coworkers. Second, his study suffered from several methodological flaws that render his results uninterpretable, even when they are considered alone.

Three basic problems with Repp's article are here addressed. The first concerns Repp's conclusions concerning geographical correlates with perception of the tritone paradox. The second concerns the presence of large

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

spectral effects in his data. The third concerns his examination of correlates between perception of the tritone paradox and the pitch range of the listener's speaking voice.

Geographical Correlates with Perception of the Tritone Paradox

PROBLEMS CONCERNING THE SUBJECT POPULATIONS

In the study by Deutsch (1991), two groups of subjects were compared. The first group consisted of 24 subjects who had grown up in California and who had spent the previous year in California. The second consisted of 12 subjects who had grown up in the south of England. No subject in the Californian group had a parent who had grown up in England, and no subject in the English group had a parent who had grown up in California. In contrast, Repp used 15 Dutch subjects, 11 subjects from the south of England, and 17 subjects from various regions of the United States. No information concerning the subjects' parental background was obtained.

It is clear that Repp's subject populations differed considerably from those of Deutsch (1991). Repp's use of Dutch subjects in this context was clearly inappropriate. Furthermore, Deutsch did not generalize her findings to all Americans. Indeed, given the recent findings by Ragozzine and Deutsch (1994) showing a regional difference within the United States in perception of the tritone paradox that depends on familial background, it is clear that such generalization cannot be made. Yet the American group used by Repp included only six Californians. So at best, assuming that the subjects in Repp's study had grown up in their "native" geographical region, and that they fulfilled the criteria concerning familial background, he would have been able to compare only 11 subjects from the south of England with 6 from Calfornia—a sample size that is too small for meaningful comparison to be made.

PROBLEMS CONCERNING THE PRESENTATION PROCEDURES

The procedures used by Repp in presenting the stimulus patterns differed in several important respects from those of Deutsch (1991) and Deutsch et al. (1990) (see also Deutsch, 1986, 1987; Deutsch, Kuyper, & Fisher, 1987). In the latter experiments, subjects were presented with tones generated under four different spectral envelopes that were spaced at half-octave intervals. Tone pairs generated under the different envelopes were presented in blocks of 12, with 5-s intervals between tones within blocks and 1-min pauses between blocks. Each block consisted of tone pairs that were generated under one of the envelopes, and the envelopes were ordered haphazardly across blocks.

In contrast, Repp presented tones generated under any one envelope in a single block of 144 tone pairs, with only 2.5 s between pairs and an extra 2.5 s following every 12th pair. This rapid and intensive presentation procedure might have induced the subjects to focus on the spectral characteristics of the tones, so that pitch class effects would, as a result, have been distorted.

Further important differences involved the spectral envelopes that were administered to the different subject populations. The Dutch subjects in Repp's study received three different spectral envelopes in counterbalanced order: one similar to that used by Shepard (1964), one similar to that used by Deutsch with the envelope peak at A_4, and one similar to that used by Deutsch with the envelope peak at $D\sharp_5$. However, this was not true of the British and American subjects: Instead of being given tone pairs generated under all three envelopes, they were given only the tone pairs generated under the A_4 and $D\sharp_5$ envelopes. Given that octave-related complexes are particularly susceptible to context effects, this difference in treatment of the subject populations is an important one. Most of the Dutch subjects received the tones generated under either the A_4 or the $D\sharp_5$ envelope immediately following the tones generated under the Shepard envelope; however, this was not true of the American subjects. Because of these differences, the results of the three-way analysis of variance (ANOVA) carried out by Repp, comparing performance of the Dutch, British, and American subject populations, cannot be meaningfully interpreted.

As a further confounding factor, Repp administered an ascending semitone scale before each of the two or three blocks of tritone pairs. This scale was administered to some of his subjects but not to others. When the scale preceded the block consisting of tones generated under either the Shepard envelope or the $D\sharp_5$ envelope, the tones ascended from D♯ to D. However, when the scale preceded the block consisting of tones generated under the A_4 envelope, the tones ascended from A to G♯ instead. Given that such tones are particularly susceptible to context effects, this preceding scale might have created differences in the data obtained from tones generated under the different envelopes.

Repp administered these ascending scales to the Dutch and British subjects (and also to the 5 Americans who were tested in the Netherlands and England) but not to the remaining 12 American subjects. Because the different subject populations were treated differently from each other, the results of the two-way comparisons between these subject populations are impossible to interpret.

DEUTSCH, DIANA, ***The Tritone Paradox and the Pitch Range of the Speaking Voice: Reply to Repp*** , Music Perception, 12:2 (1994:Winter) p.257

Spectral Effects

Repp's data contain larger differences depending on spectral envelope than do the data obtained by Deutsch and coworkers. These could have been due to a number of factors, such as differences in the subject populations or in the stimulus conditions. However, regardless of their bases, such spectral effects raise questions concerning how Repp's data should have been analyzed. In attempting to correlate his results with the subjects' vocal ranges for speech, Repp chose to focus on the data obtained only from the envelope producing the *weakest* profile (the D♯$_5$ envelope). This procedure was highly problematic, because when large spectral effects exist, it is mandatory that the data be averaged across different spectral envelopes so as to balance out effects due to the relative amplitudes of the components of the tones.

As illustration, Figure 1 presents data from an experiment in which strong effects of spectral envelope were obtained under certain conditions but not others (Deutsch, in preparation). In this experiment, perception of the tritone paradox was compared for tones containing all six octave-spaced harmonics (as is usually the case) and tones containing only three harmonics that were spaced at two-octave intervals.

Plots A and B in this illustration show the results from one subject, when presented with harmonics 1, 3, and 5 of tones generated under the envelopes centered at C_5 and F♯$_5$, respectively. These plots might lead the reader to conclude that a strong tritone paradox was obtained under these conditions, with the peak pitch classes shifting by a half-octave with the half-octave shift in the position of the spectral envelope. However, as shown in Plot C, which presents the data averaged across the two envelopes, no tritone paradox in fact emerged. Instead, under each of the envelopes, the subjects were basing their judgments on the highest amplitude components of the tones.

Plots D and E show, in contrast, the results obtained from the same subject for tones generated under the same two envelopes, when all six octave-spaced harmonics were present. Plot F shows these data averaged over the two envelopes, and it can be seen that a strong profile now emerges from the averaged data.

This example shows that when strong differences are present depending on the position of the spectral envelope, one cannot simply focus on the results obtained under a single envelope to determine the peak pitch classes. Indeed, without averaging the data over envelopes so as to counterbalance for spectral effects, one cannot even be sure that a tritone paradox has been obtained. Since Repp does not present such averaged data, it is impossible to judge whether he obtained a convincing tritone paradox in

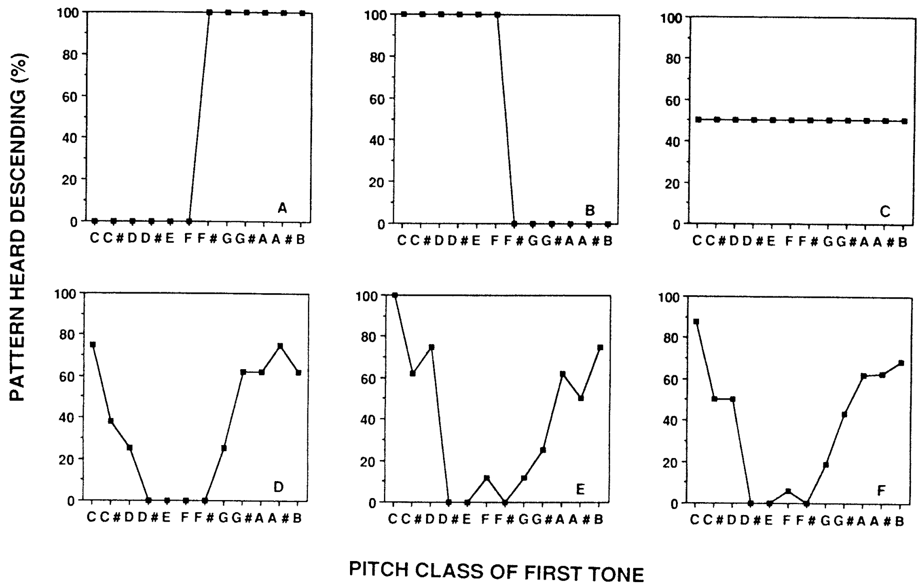

Fig. 1. The importance of averaging across spectral envelopes. Plots A and B show the data obtained for tones consisting of three harmonics that were spaced at two-octave intervals and were generated under the spectral envelopes centered at C_5 and $F\sharp_5$, respectively. Plot C shows the same data averaged across the two envelopes. Large effects of envelope position were obtained, and these cancelled each other out in the averaging process so that no tritone paradox emerged. Plots D and E show the data obtained under the same two envelopes for tones consisting of all six octave-spaced harmonics, and Plot F shows the same data averaged across these two envelopes. Here the effect of envelope position was small, and a strong tritone paradox emerged in the averaged data. Results from a single subject are here displayed.

the first place; or if he did, what the peak pitch classes for each subject would have been, had the data been averaged across spectral envelopes.

Correlates with the Pitch Range of the Listener's Speaking Voice

We now turn to the portion of Repp's paper that deals with the pitch of the listener's speaking voice. Here problems emerge both with his description of the existing literature and also with the procedures that he used.

THE EXISTING LITERATURE

Regardless of Repp's speculative arguments against the notion of a culturally acquired pitch class template, the existing literature provides strong supporting evidence in favor of such a hypothesis. A review of this literature is provided by Dolson (1994), and the following points are here emphasized:

1. The pitch ranges for individual speakers are generally close to an octave.
2. Within a linguistic community, the pitch ranges for male and female speakers are separated by roughly an octave.
3. Within a linguistic community, speakers of the same sex tend to have mean F_0s that are remarkably similar to each other. In contrast, there are large differences in group mean F_0s when different linguistic communities are compared.
4. Convincing failures to correlate F_0s with physiological measures have been reported. For example, Hollien and Jackson (1973) carried out an extensive study of 157 males, in which they examined their height, weight, head circumference, neck circumference, chest circumference (inflated and deflated), waist circumference, arm length, and leg length and took X-rays of each subject's larynx to provide indices of laryngeal size. The authors conclude, "With respect to the physical measurements obtained for these subjects, no statistically significant correlations were found between any of the voice parameters and either laryngeal size or body dimensions" (p. 120).

PROCEDURES USED TO ESTIMATE PITCH RANGE

Deutsch et al. (1990) used the following procedure to analyze the pitch ranges of their subjects' speech. A 15-min sample of spontaneous speech was recorded from each subject. From each speech sample, F_0 estimates were obtained at 4-ms intervals, so that a very large number of such estimates were obtained. These estimates were then allocated to semitone bins, and the octave band containing the largest number of estimates was thus derived.

In contrast, Repp had his subjects read 10 sentences aloud and estimated the upper and lower limits of their vocal ranges separately. The procedure that he used was problematic on several grounds. First, the pitch range for reading is known to differ from that for spontaneous speech (Dolson, 1994), so that Repp's estimates of vocal range cannot be compared with those of Deutsch et al. (1990). Second, Repp made no attempt to determine octave bands for his subjects' speaking voices, so that again, his results cannot be compared with those of Deutsch et al. (1990).

Third, as Repp himself acknowledged, the number of pitch estimates that he obtained may have been too small to obtain reliable estimates of vocal range. This concern is bolstered by the fact that Repp failed to obtain a difference between the pitch ranges of the American and British speakers. Yet as he himself acknowledged, the existing literature on American and British speech leads to the assumption that such differences should have been obtained. Fourth, despite the hypothesis advanced by Deutsch that the listener focuses on the *upper* limit of the octave band for speech to determine the highest position along the pitch class circle, Repp argues that it is preferable to examine the *lower* end of the pitch range of the subject's speech instead. This argument makes no sense at all, given that the hypothesis being examined was advanced specifically with respect to the upper limit of the pitch range of the listener's speaking voice.

In summary, Repp's findings are shown on a number of grounds to be uninterpretable in relation to those of Deutsch (1991) and Deutsch et al. (1990). (1) The geographical correlates he obtained with perception of the tritone paradox were based on subject populations that differed from those of Deutsch (1991), and his own populations were treated differently from each other. (2) Repp obtained large differences in perception of the tritone paradox that depended on the spectral envelope under which the tones were generated. Yet in attempting to correlate perception of the tritone paradox with the pitch range of the listener's speaking voice, he examined only the data obtained from tones generated under one of the envelopes that he used. (3) Repp's procedures in obtaining estimates of pitch ranges for speech were problematic in themselves and were inappropriate in the context of Deutsch's hypothesis.

References

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. The tritone paradox and central integration. (in preparation)

Deutsch, D., Kuyper, W. L., & Fisher, Y. The tritone paradox: Its presence and form of distribution in a general population. *Music Perception,* 1987, ***5,*** 79–82.

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.

Hollien, H., & Jackson, B. Normative data on the speaking fundamental frequency characteristics of young adult males. *Journal of Phonetics,* 1973, ***1,*** 117–120.

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

Repp, B. H. The tritone paradox and the pitch range of the speaking voice: A dubious connection. *Music Perception,* 1994, ***12,*** 227–255.

Shepard, R. N. Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America,* 1964, ***36,*** 2345–2353.