

Spectral-motion aftereffects and the tritone paradox among Canadian subjects

LLOYD A. DAWE

Cameron University, Lawton, Oklahoma

and

JOHN R. PLATT and EYDRA WELSH

McMaster University, Hamilton, Ontario, Canada

The effect of spectral motion on the tritone paradox was investigated by pretesting subjects residing in southwestern Ontario, Canada, on the tritone task, presenting them with a continuous ascending or descending chromatic scale created using Shepard tones, and then retesting them on the tritone task. Results indicated a negative-motion aftereffect that affected the orientation of the pitch class circle. Differential effects of perceived pitch height on the lower portion of the pitch class circle and of adaptation on the upper portion of the pitch class circle were found in the pre- and postadaptation data, respectively. The implications of this dissociation are discussed. In addition, since our subjects lived relatively close to the U.S. border, the experimental pretests allowed us to examine the hypothesis that a canonical American pitch template similar to that found among “Californian” subjects (Deutsch, 1991) is propagated by linguistic influences of media such as television and radio (Ragozzine & Deutsch, 1994). A survey of our subjects indicated that overall, the majority of time engaged in listening to the radio and watching television or movies was spent with American sources. Despite this, and despite the fact that subjects had widely varying language and cultural backgrounds, a tight distribution of peak-pitch classes was found that is indicative of a “British” pitch template (Deutsch, 1991) for every subject tested.

For more than 30 years, researchers have made use of specially designed complex harmonic tones in their investigations of different perceptual attributes associated with tonal relationships (Shepard, 1964). In these Shepard tones, adjacent partials are spaced at octave intervals and then filtered through a constant spectral envelope. Shepard tones are useful as stimuli in many psychoacoustic investigations because they ambiguate pitch height while maintaining the chroma of the tone being played. For example, repetitively playing the same 12 chromatic tones spanning an octave will create an illusion of a continuous ascending or descending pitch sequence, depending on whether one plays the chromatic series as ascending or descending, respectively (Figure 1A).

Playing *pairs* of such tones will create the perception of an ascending or descending interval on the basis of pitch proximity. If the second tone is less than half an octave (an interval called the “tritone”) above the first, one typically perceives an ascending interval, whereas if the interval is greater than the tritone, one perceives it as descending

(Shepard, 1964). What about the tritone itself? Deutsch and her colleagues (Deutsch, 1986, 1987, 1991, 1994b; Deutsch, Kuypers, & Fisher, 1987; Deutsch, North, & Ray, 1990; Ragozzine & Deutsch, 1994) have done extensive studies using such tritone pairs and have asked subjects to identify the presentations as ascending or descending. Subjects have been found to give reliable and orderly responses. If one were to envision the 12 chromatic pitch classes in a series as numbers on the face of a clock, as in Figure 1B, those tritone pairs in which the first member was in the upper half of the clock (e.g., D# through G#) would result in “descending” responses. Tritone intervals beginning with a pitch class in the lower portion of the clock (e.g., A through D) would result in “ascending” responses. Although this finding indicates that judgments are systematically related to a tone’s position within the pitch class circle, differences among subjects suggest that individuals can vary in terms of which pitch classes would be allocated to the peak positions at the top of the clock (e.g., F and F# in Figure 1B). Together, these effects constitute the tritone paradox.

Possible Underlying Neural Processes

We were interested in exploring underlying neural mechanisms and processes that may be involved in the tritone paradox in the hope that such information would provide insights into its basis. For example, many paradoxes and illusions are known to be the result of low-level neural processes (e.g., mach bands resulting from lateral inhibi-

The authors are grateful to T. Visser for her assistance in running subjects, A. Cohen for providing information regarding an unpublished study, and K. Kewish, D. Deutsch, B. Repp, and an anonymous reviewer for helpful comments. This research was supported by grants to L.A.D. and J.R.P. by the Natural Sciences and Engineering Research Council of Canada. Correspondence should be addressed to L. A. Dawe, Department of Psychology and Human Ecology, Cameron University, 2800 West Gore Blvd., Lawton, OK 73505-6377 (e-mail: lloyd@cameron.edu).

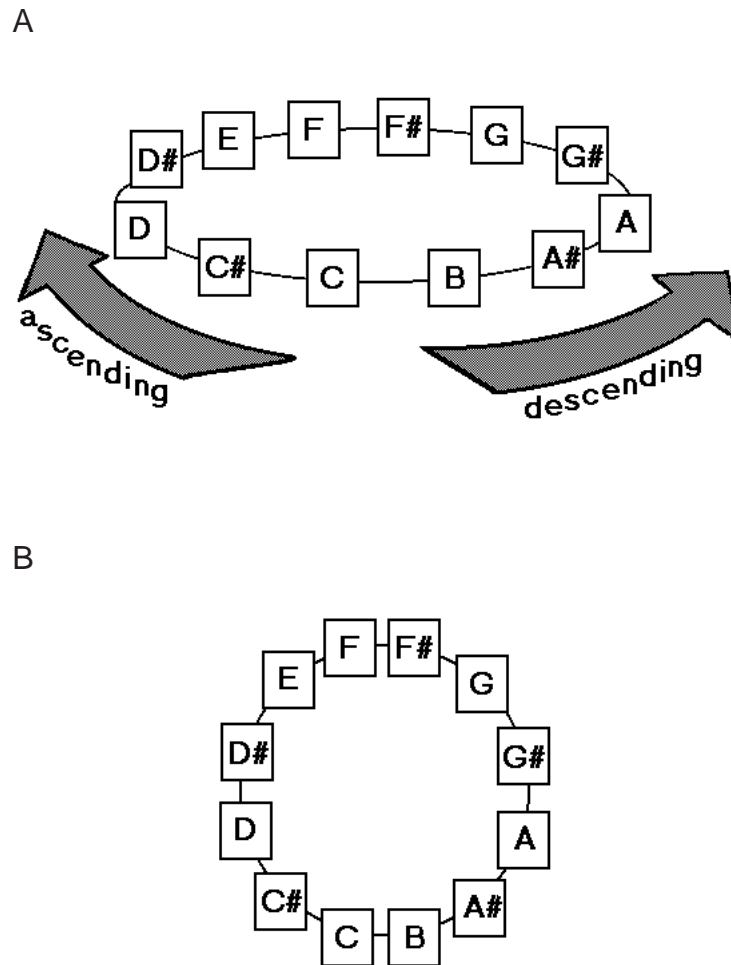


Figure 1. Using Shepard-like stimuli, judgments of tonal relationships are based on pitch class. (A) Ascending or descending spectral movement is perceived on the basis of pitch proximity. Playing notes in chromatic series (i.e., adjacent notes) can create an illusion of a continuous ascending or descending pitch spiral, depending on whether the direction is clockwise or counter-clockwise. (B) The pitch class circle can be thought of as a clock. Subjects can vary in terms of which pitch classes are placed in the peak positions at the top of the clock.

tion). However, it is unlikely that the tritone paradox, which is a rather complex phenomenon, is mediated by an exclusively peripheral mechanism. Indeed, the complexity of the stimuli and the effects of context (Repp, 1997) on the tritone paradox suggest that central processes must be involved. Further, in a 1992 study, Deutsch reported that the paradox persisted even when she presented subjects with odd-numbered octave partials in one ear while playing the even-numbered octave components in the other, indicating that structures at or above the superior olivary complex, which is where dichotic signals first interact, must be involved.

Generally, the degree of specificity demonstrated for auditory phenomena can serve as a basis for inferring a level of process within the system. For example, first-order afferent neurons in the auditory system show fre-

quency specificity, whereas those higher, at the level of the cochlear nucleus, maintain the frequency specificity and can be further classified as onset cells, offset cells, on-off cells, pausers, and choppers (Gulick, Gescheider, & Frisina, 1989). As one ascends the auditory system, lower level encodings are maintained (e.g., there is tonotopic mapping), and specificity increases. At the level of the superior olivary complex, cells have been found that respond to specific interaural time and intensity differences. By the time one reaches the auditory cortex, a large portion of cells will respond only to frequency modulations that occur in a specified direction, within a specified frequency range, and at a specified rate (Barlow & Mollon, 1982; Gulick et al., 1989; Yost, 1994). Single-cell recordings from the auditory cortex of animals in the presence of various acoustic stimuli have provided strong ev-

idence that some cells respond selectively to frequency contours, serial positions, and possibly interval distances (Weinberger & McKenna, 1988).

One cell type that exists in the auditory cortex that is of particular interest for the tritone paradox is the spectral-motion detector (Shu, Swindale, & Cynader, 1993; Whitfield & Evans, 1965). Some spectral-motion cells will respond only to ascending patterns of frequencies, whereas others will respond only to descending patterns. The experimental task employed in investigations of the tritone paradox involves a judgment of whether or not a presented pair is an ascending or descending interval; thus we hypothesized that a population of spectral-motion detectors would be involved.

Using the psychophysical procedure of adaptation, behavioral effects of spectral-movement specificity can be demonstrated. Most movement adaptation research has been conducted in the visual modality, but a few studies have been done in audition (Grantham & Wightman, 1979; Hall & Soderquist, 1982; Shu et al., 1993). Usually, a subject is presented with an adaptation stimulus that moves in a specified direction for a prolonged period of time. This presentation is believed to fatigue cells responsible for encoding movement in that direction. Subsequent tests on stationary or ambiguous patterns often result in a negative aftereffect in which movement is perceived to run in the opposite direction (e.g., the waterfall illusion or spiral movement aftereffects) due to the relatively higher spontaneous firings of cells that were not fatigued by the adaptation stimulus.

A surmountable problem with using an adaptation procedure to investigate *spectral* motion that does not exist for *spatial* motion is that at some point, one must change spectral direction. For example, a subject could not be presented with an ascending frequency glide for 2 min without either moving the frequency at a very slow rate or covering a large spectral range that extends beyond the tuning of a cell. By limiting the range and rate of spectral motion, subjects would have to be presented with iterative presentations of the same spectral motion, resulting in a rapid spectral movement in the opposite direction between iterations (see, e.g., the stimuli used by Shu et al., 1993). What is needed is an ambiguous starting and stopping point for the spectral motion to be effective and efficient as a controlled adaptation stimulus. Scales consisting of tones similar to Shepard's (1964) afford such presentations and thus could potentially serve as adaptation stimuli. Since most single-cell recording studies in animals and auditory-motion adaptation studies with humans have used simple stimuli such as pure tones (e.g., Weinberger & McKenna, 1988), the use of such complex adaptation stimuli would also serve as an extension of previous psychophysical work.

While there are undoubtedly cells of high specificity that respond to complex signals (e.g., species-specific vocal calls, Glass & Wollberg, 1983), it is unlikely that cells exist that are specifically tuned to scales made of Shepard tones. Presumably, the use of Shepard scales as

adaptation stimuli would fatigue a large number of spectral-motion-specific cells that respond to different pitch ranges. Nevertheless, such stimuli could be useful since this population of cells shares the common feature of specificity to spectral motion, which, when fatigued, may allow us to assess the contribution of spectral motion to the tritone paradox.

The pattern of responses following an adaptive context should serve as a basis from which we can infer the role of spectral-motion specificity in the tritone task. For example, one possibility is that an independent decision-making process employs spectral motion. Adaptation to an ascending sequence in such a system would result in an overall increase of *down* responses independent of the specific tritone pair being tested, whereas a descending sequence as an adaptive stimulus would result in an overall increase of *up* responses. Results supporting this hypothesis would appear as highly similar response functions for pre- and postadaptation data, with an additive shift up or down following adaptation. Alternatively, spectral-motion specificity may be directly involved with the encoding of each tritone pair. For example, if a pitch template were oriented with F and F# in the peak positions (see Figure 1B), there would be a greater descent with the tritone pairs F-B and F#-C, compared with the tritone intervals that move more horizontally within the template, such as A-D#. If the effects of adaptation are proportional to the height distance on the pitch template, effects of adaptation to ascending or descending sequences should be more pronounced for tritone pairs that are very low and high on the pitch template, respectively. Results indicative of this scenario would be postadaptation response functions that are more shallow than the preadaptation response functions. Finally, if spectral-motion-specific processes are involved in the orientation of the pitch template, adaptation should result in a reorientation of the template. Shifts to the right or left in the response function following adaptation would indicate such a relationship.

The American Media Hypothesis

Some intriguing findings have resulted from investigations of the tritone paradox. First, the limits of the octave with the largest proportion of fundamental frequencies from an individual's spontaneous speech have been found to be related to the subjectively highest or "peak" pitch classes in a tritone task (Deutsch et al., 1990). Second, there appears to be a canonical distribution of peak-pitch classes for subjects in the United States (Ragozzine & Deutsch, 1994). Third, subjects from southern England have been found to have peak-pitch class distributions that are essentially the opposite of those found for most American subjects (Figure 2).

Deutsch and her colleagues (Deutsch, 1991; Deutsch et al., 1990) have offered a developmental learning explanation for these findings in which the pitch class circle is acquired by exposure to speech sounds within one's linguistic community. The limits of the octave band for speech appear to be represented on this mental template as

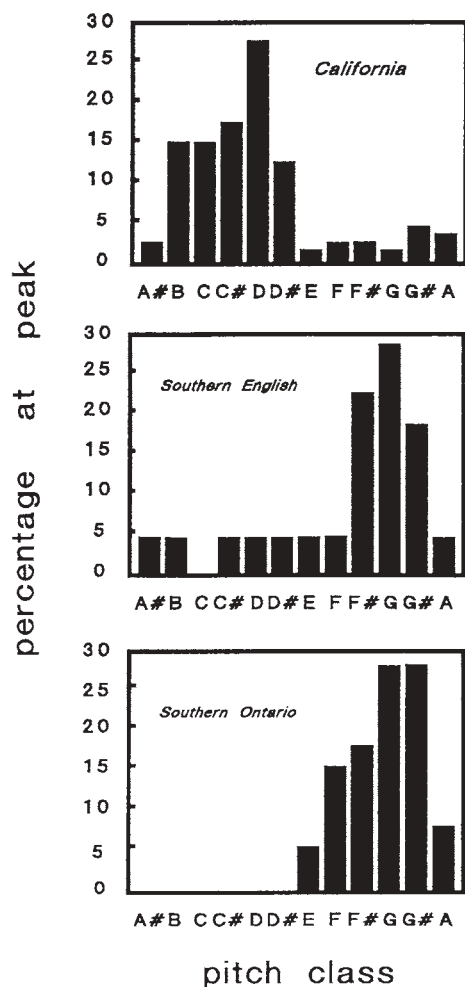


Figure 2. Distributions of peak-pitch classes for the present study's preadaptation data (bottom graph), California (top), and southern English (middle) subject populations. In all cases, the data are averaged over tones generated under the four spectral envelopes. Note—The top two graphs are reprinted from "The Tritone Paradox: An Influence of Language on Music Perception," by D. Deutsch, 1991, *Music Perception*, 8, p. 343. Copyright 1991 by the Regents of the University of California. Reprinted with permission.

the peak-pitch classes. Although this explanation has been met with some criticism (Repp, 1994; see also Deutsch, 1994a), no alternative explanation has been proposed.

In 1994, Ragozzine and Deutsch explored the possibility of a regional difference in perception of the tritone paradox within the United States. They ran two groups of subjects who resided in Ohio—one called "local" on the basis that both parents had grown up within the same area, and the other called "alien" on the basis that at least one parent had been raised within another area of the United States. They found that the alien group and approximately half of the local group had a peak-pitch class distribution similar to that found for California subjects (Figure 2).

The other half of the local group had peaks within the opposite half of the pitch class circle. In general, profiles for the local group were much less pronounced than were those for the alien group. These data, in combination with informal observations from a large number of subjects across the United States indicating a typical profile similar to the California responses (Ragozzine & Deutsch, 1994), led Ragozzine and Deutsch to hypothesize that a canonical manner of hearing the tritone stimuli exists within the United States and that this may be propagated by the linguistic influences of media such as radio and television. The fact that the local subjects could be split into two groups on the basis of their peak-pitch response distributions may reflect a gradual change from a local, regional orientation of the pitch class circle to the canonical one.

Our experiment was run in a southern portion of Ontario within a 1.5-h drive from the American border, which afforded us the opportunity to test these hypotheses. Our subjects had easy access to American media such as radio and television, as well as theater and concerts, which, if Deutsch's hypothesis is correct, could result in less pronounced profiles and/or a predominantly American peak-pitch class template. Alternatively, the presumably closer linguistic ties Canada shares with England may result in subjects having a predominantly British peak-pitch class profile. To our knowledge, there has been no direct comparison between British and Canadian speech characteristics in terms of their predominant speech octave bands. However, British spelling and pronunciation are common in Canada, so it would not be surprising if the two countries shared linguistic characteristics.

We anticipated that other factors might play a role in the response profiles of our subjects. Canada is often characterized as a large, multicultural nation in which many ethnic groups maintain much of their cultural and linguistic identities. For example, just within the metro Toronto area, more than 50 languages are spoken and many ethnic communities can be found. In addition, Canada has a relatively high acceptance rate for immigrants and refugees (e.g., 70% compared with 17% for the United States for acceptance of refugees). Finally, there are widely diverse regional accents and dialects across Canada. These factors led us to expect a heterogeneous subject pool in which, regardless of the predominant peak-pitch class profile found (i.e., Californian or British), the peak-pitch class would quite likely be variable.

METHOD

Subjects

Twenty introductory psychology students at McMaster University participated in the experiment. They ranged from 19 to 32 years of age with a median of 19 years. Thirteen subjects were female and 7 were male. Eleven subjects had more than 4 years of formal musical training and 9 were nonmusicians. Subjects received course credit for participating in the experiment and additional course credit for completing the survey, which was conducted several months subsequent to the experiment.

Equipment

Each subject was tested individually in a small AEC sound-attenuating chamber. All auditory stimuli were generated digitally in real time with 16-bit precision by a Symbolic Sound Corporation Kyma Sound Design Workstation running at a sample rate of 22.05 kHz. The output of this system was low-pass filtered with a cutoff frequency of 8 kHz and passed through a NAD Model 3020e stereo amplifier before being presented to the subject by a matched pair of THD-39 earphones. The Kyma system was in turn programmed and controlled by a 486/66 computer, which also recorded subjects' responses to the stimuli. Subjects responded by clicking a Logitech mouse on the appropriate button of a virtual control panel displayed on a Mitsubishi Diamond Scan 20LP video monitor. All auditory stimuli were set to a level of approximately 70 dB SPL with the aid of a Bruel and Kjaer Type 2231 Integrating Sound Level Meter and Type 4152 artificial ear.

Stimuli

Preadaptation tritone pairs. To facilitate comparison of our findings with those reported in the literature, we used the stimuli generation procedures reported by Deutsch et al. (1987) and Ragozzine and Deutsch (1994). Each complex tone consisted of six sinusoids spaced at octave intervals. The relative amplitudes of partials were determined by filtering them through a constant bell-shaped spectral envelope described by the following equation:

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

The amplitude of each frequency component, $A(f)$, was a function of β , the ratio of adjacent sine waves, which was set at 2; γ , the number of sine waves within the complex, which was set at 6; and f_{\min} , the minimum frequency for which a nonzero amplitude was assigned. Following the procedure of Deutsch et al. (1987), four envelopes at half-octave spacings were employed to balance possible effects of relative amplitudes of the sinusoidal components on judgments (see Repp, 1997). The peaks of these spectral envelopes were 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). The 12 tritone pairs (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) were generated under each of the four spectral envelopes. Since the chromatic series divides an octave span into 12 equal steps on a log-transformed frequency axis, the frequency of each chromatic tone was determined using the following equation:

$$f_n = f_{\min} (2^{n/12}),$$

where f_n is the frequency of the chromatic tone that is n semitones away from f_{\min} —the minimum frequency of the spectral envelope. Each member of the tritone pair sounded in succession for 0.5 sec with 10-msec rise and fall times and no silent interval between the members of a pair. A subject's directional response to each tritone pair initiated the next trial.

Adaptation scales. The same method used to create the complex tones in the tritone pairs was used to create two sets of adaptation stimuli, one ascending and the other descending in chromatic series, under each of the four spectral envelopes described above. Figure 1A illustrates the two sets of adaptation stimuli in which the chromatic tones were presented in a descending (i.e., counterclockwise) or ascending (i.e., clockwise) direction. Each tone had a duration of 0.5 sec with 10-msec rise and fall times and no silent interval between tones. Each adaptation scale was repeated eight times in succession for a total duration of 48 sec. A subject's response as to the perceived direction of the adaptation scale served as a trigger to initiate the postadaptation trials.

Postadaptation stimuli. Postadaptation tritone pairs were identical to those in the preadaptation pairs, except that each pair was

preceded by 3 sec (i.e., the first 6 notes) of the adaptation scale. Thus for ascending adaptation conditions, the 6-note sequence was C, C#, D, D#, E, F for the spectral envelopes centered on C, and F#, G, G#, A, A#, B for the spectral envelopes centered on F#. The descending adaptation conditions had a sequence of B, A#, A, G#, G, F# for the spectral envelopes centered on C, and F, E, D#, D, C#, C for the spectral envelopes centered on F#. The chromatic sequences served as a top-up to maintain a relatively constant level of adaptation over the test period. A 1-sec silent interval separated the scale from the tritone pair, and a subject's directional response to each pair initiated the next trial.

Questionnaire characteristics. An extensive questionnaire was designed to collect information on each subject's culture and family history; languages spoken within the family; educational history; musical background; music preferences; and time spent listening to music and radio, watching television, and going to theaters, movies, or concerts. The questionnaire was designed to give a broad picture of any encounters over the subject's lifetime with different languages.

Procedure

Subjects were randomly and equally assigned to one of two adaptation groups (ascending or descending scales). Each subject participated in four experimental phases with all stimuli (tritone pairs and adaptation stimuli) within each phase generated under a specific spectral envelope. The envelopes were always run in the order C_4 , $F\#_4$, C_5 , $F\#_5$. Each phase consisted of 48 preadaptation trials (12 tritone pairs presented in each of four different random orders with the restriction that the same pitch classes did not occur in any two consecutive pairs), followed by 48 sec of an adaptation scale (either ascending or descending for a particular subject) and 48 postadaptation trials (same as preadaptation trials except for a 3-sec adaptation scale preceding each tritone pair). The four phases were completed within one experimental session with a short break between phases. Responses were registered to the computer via a virtual control panel presented on the monitor. By moving a computer mouse, the subject could place the cursor into one of two boxes, labeled "up" and "down." Clicking the mouse registered the response. Subjects were told that they would hear a pair of tones on each trial and that they were to click on the *up* button if the second tone sounded higher in pitch than the first tone. If the second tone sounded lower in pitch than the first, they were to click on *down*. The instructions were present on the virtual control panel throughout the duration of the experiment. Subjects completed the questionnaire as a follow-up to the experiment.

RESULTS AND DISCUSSION

Questionnaire Results

Eighteen of the 20 participants completed and returned the questionnaires. As can be seen from Table 1, a wide variety of ethnic backgrounds were represented in our subject pool. Only 67% of respondents were born in Canada, and only 39% had at least one parent born in Canada. From a developmental learning perspective, our subject pool had access to a wide variety of linguistic characteristics while growing up. Approximately 56% had received formal education in more than one language, 83% were raised in a multilingual home environment, and 78% considered themselves as multilingual. It is also important to note that 22% of the respondents did not identify English as their principal language. To the extent that the linguistic characteristics differ—and, in particular, the limits of the speech octave bands of these languages differ—we expected to find a highly variable peak-pitch class distribution.

Table 1
Countries of Family Members' Birth and Languages Spoken

Subject	Father	Mother	Languages Spoken at Home*
1 Canada	Canada	Canada	<u>French, English</u>
2 Canada	Scotland	Scotland	<u>English, French</u>
3 Canada	Scotland	Scotland	<u>English, French</u>
4 Canada	Canada	Wales	<u>English, Signed English, French</u>
5 Canada	Germany	Canada	<u>English</u>
6 Canada	England	Canada	<u>English, French</u>
7 Hong Kong	Hong Kong	Hong Kong	<u>Cantonese, English, Mandarin</u>
8 Canada	Italy	Italy	<u>Italian, English, French</u>
9 Canada	Hong Kong	England	<u>English, French</u>
10 Romania	Romania	Romania	<u>Romanian, English, Hungarian</u>
11 Trinidad	Trinidad	Trinidad	<u>English</u>
12 Vietnam	China	China	<u>English, Chinese, French, Vietnamese</u>
13 Pakistan	India	India	<u>English, Urdu</u>
14 United States	Canada	United States	<u>English, French</u>
15 Canada	Italy	Italy	<u>English, Italian, French</u>
16 Canada	Canada	Canada	<u>English, French</u>
17 Canada	Netherlands	Canada	<u>English</u>
18 Canada	Trinidad	Grenada	<u>English, French</u>

* Languages spoken at home are listed ordinally from most to least frequent. Those spoken by the subject are underlined, and languages of schooling are in boldface.

Recently Deutsch (1996) reported that a subject's perception of the tritone paradox is highly influenced by his/her mother's native language. In this respect, it is important to note that in the present study, there is greater heterogeneity among the mothers' countries of birth than among the subjects' countries of birth, a factor which should also contribute to a wide distribution of peak-pitch classes.

We were particularly interested in the profiles for subjects born and raised in other countries and for whom British English was not their principal language (e.g., Subjects 7, 10, 12, and 14). In particular, Subject 14, whose mother was born in the United States, spent the first 12 years of his life in the United States; he was born and lived for 1 year in Virginia, then 2 years in Michigan, followed by 1 year back in Virginia, and 8 years in Pennsylvania. If a canonical pitch template exists within the United States that one adopts via developmental learning, we would expect Subject 14 to have a profile similar to Californian subjects.

In respect to pastime activities, there was high intra- and intersubject variability in the number of hours spent in the various activities. Overall, subjects indicated that they spent an estimated 23.5 h per week on the average listening to the radio, watching television, or attending a movie/watching a video. However, the estimated time spent engaged in these activities ranged from 5 to 78 h per week across subjects. Of the three activities represented, there was more time spent watching television ($M = 10.9$ h, $SD = 7.4$), followed by listening to the radio ($M = 9.4$ h, $SD = 10.22$), and finally watching videos or movies ($M = 3.1$ h, $SD = 1.8$). Of particular concern with respect to this study was the estimated proportion of time spent with American sources. Approximately 83% of the movies seen by the respondents were both produced in the United States and consisted of American content and setting; 73% of the time spent watching television was allocated to an American channel; and 36% of the time spent

listening to the radio was with an American station. Overall, of the total time engaged with these various forms of entertainment, the most (60% on the average) was with American content. On the basis of this information and assuming that Ragozzine and Deutsch's (1994) hypothesis regarding the role of media in propagating a canonical pitch template is correct, one would predict that our subjects would have peak-pitch class distributions similar to those of the California subject pool.

Preadaptation Data

For each of the 20 subjects, the proportions of downward interval judgments were averaged across the four spectral envelopes and then plotted as a function of the first tone in each pair. For each subject, the pitch class circle was bisected to maximize the difference between the upper and lower halves. Orienting the pitch class circle so that the bisection line was horizontal, we were able to identify the two pitch classes at the top of the circle as the peak-pitch classes. The resulting distribution of peak-pitch classes is shown in Figure 2 along with the distributions reported by Deutsch (1991).

It is clear that the subjects participating in this study had peak-pitch classes similar to those found for the southern English subjects in Deutsch's (1991) study. Indeed, every subject had peak-pitch classes between E and A, including Subject 14, resulting in a tighter profile than that found by Deutsch even though this subject pool was highly variable in terms of its linguistic histories and the proportion of time spent listening to the American media outlets.

This result is difficult to reconcile with Ragozzine and Deutsch's (1994) hypothesis. One might argue that since the questionnaire indicated the time spent engaged with various media and entertainment outlets at the time of survey completion, the results do not speak to the developmental learning hypothesis. Perhaps a critical period exists for the adoption of a particular orientation of the pitch

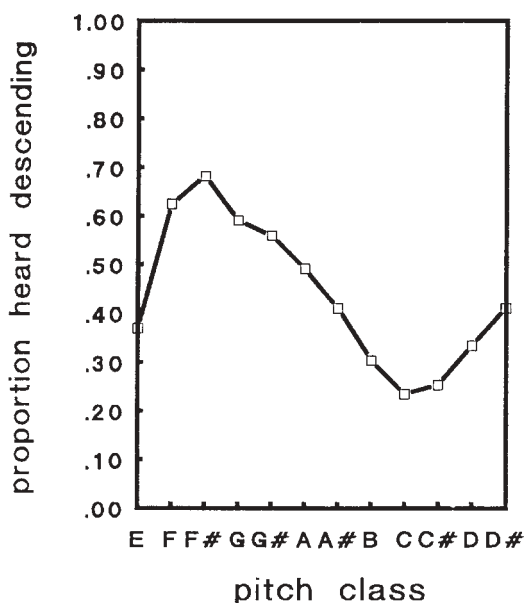


Figure 3. Proportion of judgments that a tone pair forms a descending pattern, plotted as a function of the first pitch class of each tritone pair, averaged over four spectral envelopes and 20 subjects. Individual profiles do not differ substantially from the average profile.

class circle, in which case the time spent engaged with various media at an earlier point in life would be more pertinent. However, given the highly variable linguistic histories of our subjects, one would still expect a variable distribution of peak-pitch classes, and such an explanation would not explain why our American-born and raised subject had southern English peak-pitch classes.

With respect to the relative tightness of the peak-pitch class distribution, one might interpret the result as indicative of a relatively rapid adoption of a "Canadian-British" template when individuals move to Canada. However, there is evidence that the southern English profiles found in this study are not characteristic of all Canadians. In an unpublished study conducted in the Maritimes (MacKinnon, 1993), Canadian-born and raised subjects from eastern Canada had primarily Californian profiles. Further, in subsequent experimental work we found 2 subjects with Californian profiles who were born and raised in southern Ontario. What determines a Californian or southern English profile across individuals remains a mystery.

The proportion of *down* responses as a function of the first tone of the tritone pairs averaged over subjects is shown in Figure 3. A 2 (adaptation conditions) \times 12 (tritone pairs) mixed analysis of variance (ANOVA) conducted on the proportion of down responses indicated no evidence of a main effect of adaptation group [$F(1,18) = 4.2$, $MS_e = .047$, $p > .05$] or of an interaction between the adaptation condition and tritone pairs [$F(11,198) = 1.109$, $MS_e = .024$, $p > .05$]. There was, however, a significant main effect of tritone pairs [$F(11,198) = 18.738$, $MS_e =$

.024, $p < .001$], driven by the high intersubject agreement in orientation of pitch class circles, resulting in a relatively small error variance. The results illustrate the effect of the pitch class circle on tritone intervals and are typical of the pattern found by Deutsch and her colleagues.

In 1994, Deutsch reported a further geographical correlate of Californian and southern English response profiles (Deutsch, 1994b). She found support for her hypothesis that English subjects would produce stronger profiles under the higher envelopes and weaker profiles for tones presented under the lower envelopes. Californian subjects were hypothesized to show opposite effects in which stronger profiles would be found for tones generated under the lower

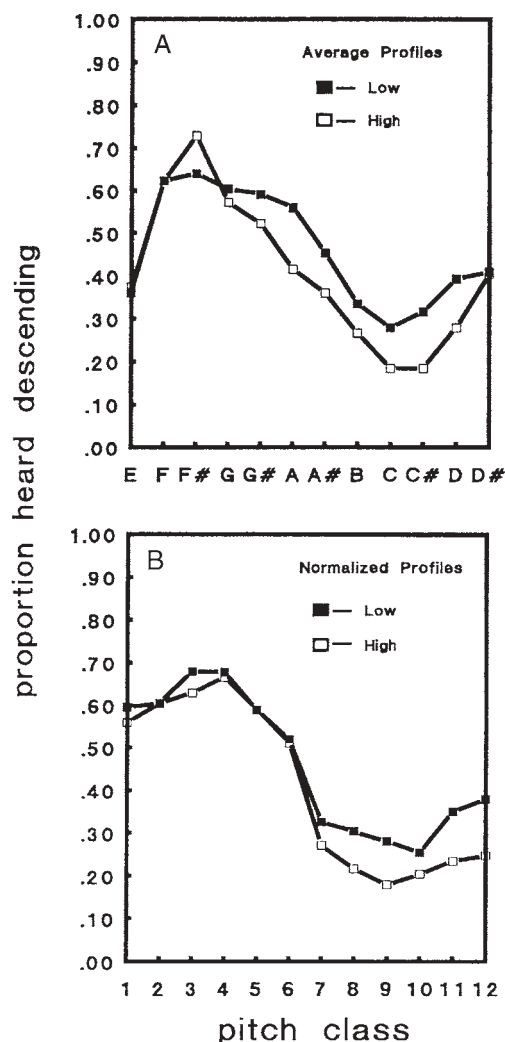


Figure 4. Proportion of judgments that a tone pair forms a descending pattern, plotted as a function of the first pitch class of each tritone pair, averaged over 20 subjects. (A) This graph represents the averaged profile for tones generated under the high (open squares) and low (filled squares) spectral envelopes. (B) This graph represents the same data first normalized across subjects.

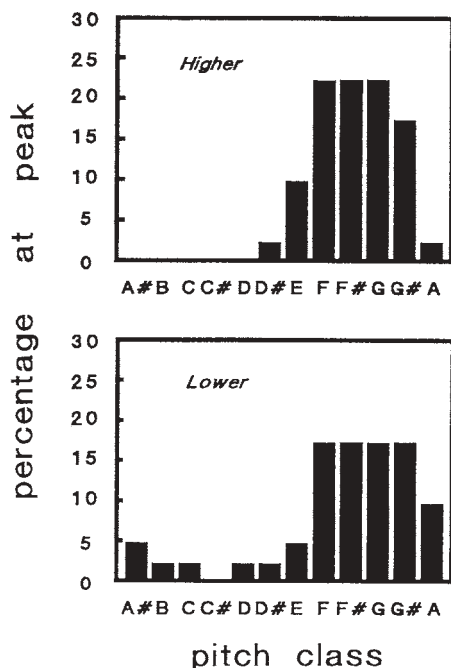


Figure 5. Distributions of peak-pitch classes among subjects for tones generated under the higher spectral envelopes (upper graph) and lower spectral envelopes (lower graph).

envelopes than those generated under the higher ones. These hypotheses were based on reported speech characteristics of a low pitch range for Californian speech (Hanley, Snidecor, & Ringel, 1966) and high pitch excursions for British English speech (Collier, 1991; Willems, Collier, & 't Hart, 1988).

Following the procedure employed by Deutsch (1994b), we divided the preadaptation data on the basis of tones generated under the high (C_5 and $F\#_5$) and low (C_4 and $F\#_4$) spectral envelopes. The top panel of Figure 4 illustrates the results of this analysis. Figure 4B shows the same data after these had been normalized for each subject for tones presented under the higher and lower envelopes (Deutsch, 1994b) to control for possible artifacts that might have arisen due to the averaging process: Although the peak-pitch class distribution was relatively tight, there was still some variability in the orientation of the pitch templates (i.e., ± 3 semitones about the typical peak-pitch classes of G, G#), which could have resulted in an artifact in the average profile. To normalize data, two separate pitch class circles for each subject, one for each spectral-envelope condition (high and low), were bisected to maximize the differences between the upper and lower halves. Orienting the pitch class circle so that the bisection line was horizontal, numbers were assigned beginning with the left-most pitch of the upper half of the circle and moving in a clockwise direction. The data were then averaged across subjects for each numerical position. The effects of envelope are evident in both representations.

As reported by Deutsch (1994b), we found that tones generated under the higher envelope resulted in a more pronounced profile, whereas tones generated under the lower envelope resulted in weaker response profiles. One might be inclined to view this pattern of results as a procedural artifact. Because we presented the four envelope phases in a systematic fashion moving from low to high, the more pronounced profile for the higher envelopes could be a mere practice effect. However, three pieces of evidence support our contention that the results are not an artifact.

First, an analysis of profiles generated under each of the four spectral envelopes indicated a highly similar pattern of results for the first two envelope conditions (C_4 and $F\#_4$) that was different from the highly similar pattern of results for the last two envelope conditions (C_5 and $F\#_5$). Thus instead of a progressively more pronounced profile under each subsequent spectral-envelope condition (which one would expect from practice effects), a qualitative shift was observed between the first and last two envelope conditions. Second, the similarity between Deutsch's data (see Figure 4 of Deutsch, 1994b) and ours is quite striking. In both Deutsch's southern English data and ours, the more pronounced profile for tones generated under the higher envelope was driven by a lower proportion of *down* responses for the lower section of the pitch class circle (positions 7–12 of Figure 4B), with no apparent change in responses for the pairs that originate in the upper half of the circle. This was substantiated by post hoc analyses. Using the normalized data for the upper half of the pitch class circle, an ANOVA comparing the high- and low-envelope conditions was conducted. No significant difference due to envelope condition was found [$F(1,19) = 0.273$, $MS_e = 0.06927$, $p = .60$]. A similar test conducted on the normalized data for the lower half of the circle did result in a significant effect of envelope condition [$F(1,19) = 10.1066$, $MS_e = 0.04862$, $p < .005$]. If the more pronounced profile for higher spectral envelopes was due to practice effects, why were there effects for only the lower portion of the pitch class circle? Third, the California data reported by Deutsch showed the opposite results, with a greater proportion of *down* responses for the upper section of the pitch class circle for tones generated under the lower envelope and no changes in responses for pairs that originated in the lower half of the circle, and they were run under the same procedures as her southern English subjects.

We believe that our results replicate Deutsch's (1994b) in this respect and that the effects suggest a convergent-octave or pitch-height referent that is located in the upper half of the pitch template for subjects with the southern English mode of responding, and in the lower half of the pitch template for subjects with the Californian mode of responding. This would place the referent around the chromas $F\#$ – $G\#$ and would result in little or no effect of octave register for tones with these chromas. The differential placement of the referent relative to the template for subjects with Californian and southern English modes of responding represents another geographical correlate of

the tritone paradox and suggests that different processes may be involved for the upper and lower halves of the pitch class circle. If the dissociation is veridical, as the similarity between Deutsch's (1994b) and our results suggests, it may be possible to find other differential effects for the upper and lower halves of the pitch class circle as a result of experimental manipulations.

Deutsch (1994b) also found evidence in support of her hypothesis that the peak-pitch classes would be clustered more for the more pronounced envelope condition than for the weaker envelope condition. If a similar effect were to be found here, one would predict a more clustered profile of peak-pitch classes generated under the higher spectral envelope. The profiles for the two envelope conditions shown in Figure 5 support this hypothesis.

As a final analysis of the preadaptation data, we examined the data for possible effects of spectral envelope. In an investigation of spectral effects, Deutsch (1987) reported nearly invariant results across envelopes. Repp (1994), however, found variable results in a larger subject pool. Approximately 30% of his subjects showed a reversal (i.e., a shift in the response functions of 5–6 semitones) in their pattern of results for spectral envelopes centered on A_4 and $D\#_5$, whereas 35% showed invariant patterns and the rest showed intermediate patterns. Repp's results indicate that subjects can perceive the tritone stimuli in terms of either spectral characteristics or pitch class. For this analysis, we noted the peak-pitch classes for each subject under each of the four envelope conditions, as well as the number of semitones each peak-pitch class pair was from the typical pair of G and G# (see Figure 2C). Only 2 subjects showed evidence of a consistent reversal of response patterns when shifting from a spectral envelope centered on F# to one centered on C. Both subjects responded with peak-pitch classes that were the same as the spectral-envelope peaks (i.e., C–C# for envelopes centered on C and F#–G for envelopes centered on F#). Further, both subjects had more pronounced profiles for envelopes centered on F# than on C within their respective octaves (i.e., $F\#_4$ was more pronounced than C_4 , and $F\#_5$ was more pronounced than C_5). Five other subjects showed a reversal on only one of the four spectral-envelope conditions and were consistent on the remaining three. Overall, 71 out of 80 conditions (four envelopes \times 20 subjects) had peak pitch classes within 3 semitones of the G–G# peak-pitch class pair.

Postadaptation

For each subject, the proportions of downward interval judgments were averaged across the four spectral envelopes for postadaptation data. A 2 (adaptation conditions) \times 12 (pitch class) ANOVA conducted on the data indicated a significant effect of pitch class [$F(11,198) = 5.24798$, $MS_e = .05603$, $p < .001$] and a significant interaction between pitch class and adaptation condition [$F(11,198) = 4.8811$, $MS_e = .05603$, $p < .001$]. The postadaptation data were then compared to preadaptation data for each subject. Figure 6A illustrates the results of this comparison aver-

aged over subjects in the ascending spectral-adaptation condition. Results averaged across subjects in the descending spectral-adaptation condition are shown in Figure 6C. Figures 6B and 6D show the pre- and postadaptation data after these were normalized for each subject on the basis of the preadaptation responses and then averaged for the ascending and descending conditions, respectively.

There are three notable effects of adaptation evident in the representations of Figure 6. First, there is an absolute difference in proportion of *down* responses between pre- and postadaptation as a function of pitch class. This effect is easier to see in the normalized data representations of Figures 6B and 6D. A larger effect for both forms of adaptation is found for the upper half of the pitch class circle (positions 1–6 of the normalized representations) than for the lower half. This finding supports an earlier hypothesis that the upper and lower portions of the pitch class circle are mediated by different processes. Spectral-envelope effects (high vs. low) were apparent in the lower portion of the circle and effects of spectral-motion adaptation were apparent in the upper half of the circle.

Second, a shift in the orientation of the pitch class circle was found. The shift is more apparent as a result of descending adaptation (see Figure 6C) than of ascending adaptation (see Figure 6A) and is represented by a shift to the right or left in the response function. It is important to note two things regarding this effect. First, the graphical representations in Figure 6 were affected by artifacts due to the averaging of the profiles across subjects. Some individuals had very pronounced profiles and effects of adaptation, whereas others had shallow functions. The average profiles presented as Figures 6A and 6C place more weight on the subjects with more pronounced profiles rather than equal weighting of each subject's responses. Second, it is not possible to ascertain the direction of peak-pitch class shift since the pitch classes form a circle (e.g., a shift of 4 semitones upward could actually be a shift downward of 8 semitones). With these problems in mind, we noted the number of semitones, moving in a clockwise direction, that the peak-pitch classes shifted from preadaptation to postadaptation for each individual subject in order to assess adaptation shifts of the pitch class circle. The average shift for the ascending adaptation condition (either +4.4 semitones or –7.6 semitones) was significantly different from zero [$t(9) = 2.96$, $p < .01$]. The average shift for the descending condition (either +4.5 semitones or –7.5 semitones) was also significantly different from zero [$t(9) = 3.55$, $p < .01$]. This finding indicates that perception of the tritone interval as ascending or descending is dependent on spectral motion. Specifically, the orientation of the pitch class circle appears to be mediated in part by ascending and descending spectral-motion detectors.

Third, the direction of the difference in the proportion of *down* responses between pre- and postadaptation as a function of the type of adaptation should be noted. The results are interpretable in terms of a negative aftereffect.

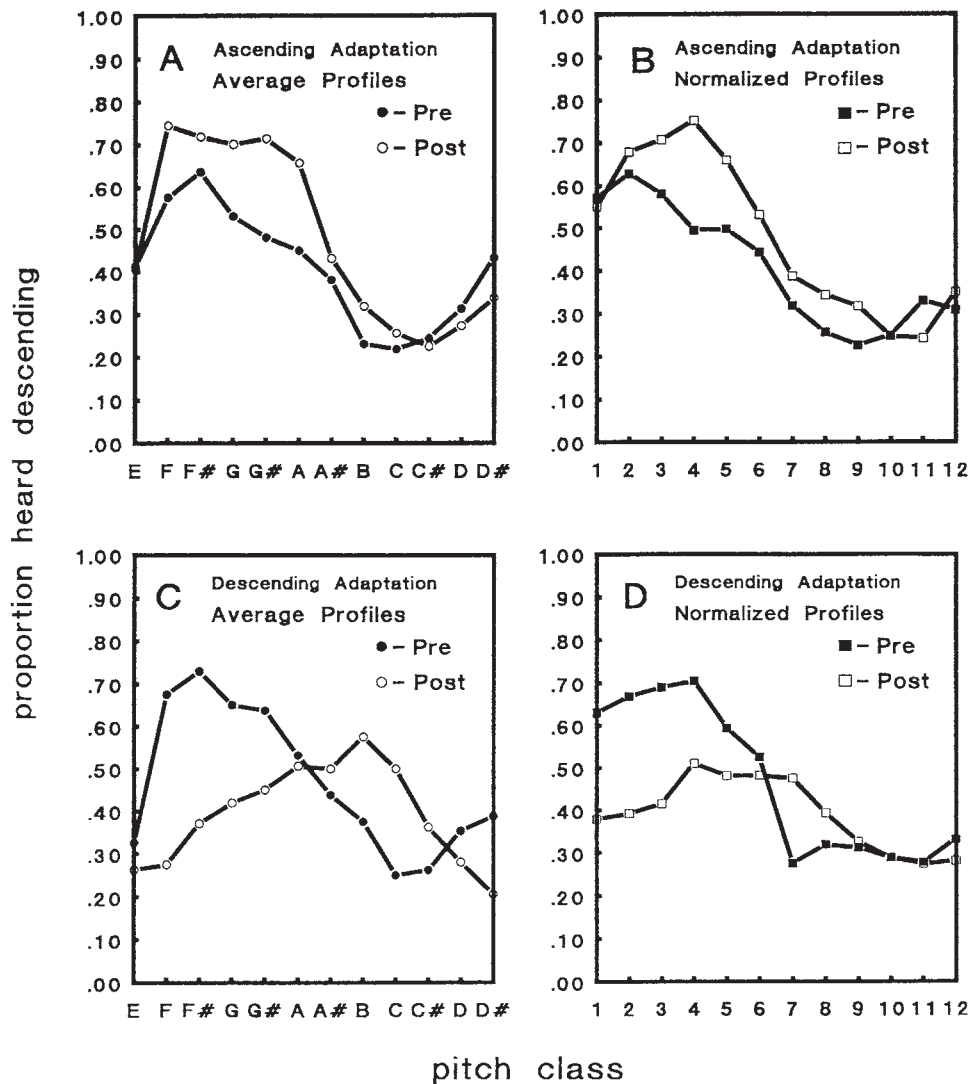


Figure 6. Proportion of judgments that a tone pair forms a descending pattern plotted as a function of the first pitch class of each tritone pair averaged over 20 subjects. The two upper graphs (A and B) represent preadaptation (filled symbols) and postadaptation (open symbols) responses for the ascending adaptation condition. The two lower graphs (C and D) represent preadaptation (filled symbols) and postadaptation (open symbols) responses for the descending condition. The graphs on the right (B and D) represent the data first normalized across subjects on the basis of preadaptation responses, and the graphs on the left (A and C) represent the average profiles.

Ascending spectral adaptation resulted in an increase in descending judgments, while descending spectral adaptation resulted in an increase in ascending judgments. This is the typical result of an adaptation paradigm. Most importantly, the effects of adaptation were not uniform across pitch classes, which would be expected if spectral-motion were involved in an independent decision-making process. The effects of adaptation as a function of pitch class provide additional evidence that spectral-motion specificity is directly involved in the pitch class “template.”

Contextual Interpretations

It must be noted that it is possible to interpret the negative aftereffect in terms of contextual contrast as opposed

to physiological adaptation. It is often difficult with adaptation experiments to separate possible contextual interpretations from adaptation interpretations, particularly when one is dealing with a central process. In this case, the presence of the 6-note top-up adaptation sequence 1 sec before each trial might have served as a context within which the tritone pair was judged. Repp (personal communication, July 12, 1996) has noted through informal explorations that the final tone in a random sequence of the 12 pitch classes can bias the perception of the subsequent tritone pair. In this regard, it is important to note that the final tone for the adaptation top-up was a correlate of the spectral envelope, and not the adaptation condition. Specifically, the ascending sequence for spectral envelopes cen-

tered on C played in chromatic series from C up to F, and the descending sequence played in chromatic series from B down to F#. Similarly, the final tones of the ascending and descending sequences for the spectral envelopes centered on F# were B and C, respectively. If the final tones of these presentations served as a contextual cue, one would predict a reversal of the response profiles across the spectral envelopes centered on C and F# and a similar pattern of results across adaptation conditions. In contrast, Figure 6 shows a clear difference between adaptation conditions, and an examination of the peak-pitch classes for the postadaptation response profiles for each envelope and subject indicated no differences between spectral-envelope conditions centered on F# and C. These findings indicate that any contextual interpretation of these effects should not be based on the final note of the sequence alone.

A contextual interpretation based on the specific notes employed within the entire sequence also has its problems. A large effect of the adaptation sequence was found for the upper half of the pitch class circle (see Figures 6B and 6D) and a minimal effect was found for the lower half. Why would a contextual effect not be evident for tones in the lower half of the circle? And, since the same pattern of responses was found for ascending note sequences beginning on F# and C, and descending sequences beginning on F and B, why were the results the same when different note contexts were employed? We believe that the only contextual interpretation viable for our data is a relative one in which the adaptive context is identified as ascending or descending without concern for the actual notes employed. In our opinion, the interpretation of these results in terms of adaptation effects with a dissociation between the upper and lower half of the circle is parsimonious in that it is in line with the dissociation evident in the preadaptation data. This position is tentative and further research is necessary to determine the level at which the effects are operating.

SUMMARY AND CONCLUSIONS

In summary, the results of our study did not support Ragozzine and Deutsch's (1994) hypothesis that a canonical template is adopted via the linguistic influence of American media. Our subjects indicated that a large portion of their time was spent watching American television and engaged in American entertainment; nevertheless, a southern English profile was found. The varied linguistic backgrounds of our subjects and relatively tight peak-pitch class distribution may indicate that the orientation of the pitch class circle is malleable. Subjects born and raised in a variety of linguistic environments may adopt the template of a regional majority when they relocate. If this hypothesis is true, the regional template in southern Ontario appears to be predominantly British, but not exclusively so. Explanations of the tritone paradox will have

to account for the highly similar profiles found on a regional basis despite varied linguistic backgrounds.

Our use of Shepard scales represents a departure from the traditional use of simple sinusoidal stimuli in a motion-adaptation study. The use of such complex stimuli afforded us the opportunity to investigate the role of spectral-motion specificity in the tritone paradox. Evidence was found indicating that spectral-motion-specific processes are directly involved in the paradox. Evidence was also found of a dissociation between the upper and lower halves of the circle. High versus low spectral-envelope effects were noticeable in the lower portion of the pitch class circle, while effects of spectral motion were most noticeable in the upper half of the pitch class circle. A reversal of the position of these two effects for subjects showing a Californian peak-pitch profile would serve as strong evidence that the dissociation is veridical. It is clear that the underlying processes are much more complex than the tritone task of ascending or descending judgments implies.

We believe that much more research is necessary to reach an adequate explanation of the tritone paradox, and that one fruitful area of research may be a further description of the underlying processes involved. To this end, Shepard tones and the spectral-motion adaptation paradigm may offer distinct advantages because they allow for the independent manipulation of range and rate of spectral-motion without the confound of a rapid spectral shift. Future studies should concentrate on determining the level at which these effects are operating and the independent manipulation of range and rate of spectral motion, as well as the interval distance between adjacent partials of the Shepard tones, in an effort to gain a fuller description of underlying processes and a greater understanding of this perplexing paradox.

REFERENCES

- BARLOW, H. B., & MOLLON, J. D. (1982). *The senses*. Cambridge: Cambridge University Press.
- COLLIER, R. (1991). Multi-language intonation synthesis. *Journal of Phonetics*, *19*, 61-73.
- DEUTSCH, D. (1986). A musical paradox. *Music Perception*, *3*, 275-280.
- DEUTSCH, D. (1987). The tritone paradox: Effects of spectral variables. *Perception & Psychophysics*, *41*, 563-575.
- 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. (1994a). The tritone paradox and the pitch range of the speaking voice: Reply to Repp. *Music Perception*, *12*, 257-263.
- DEUTSCH, D. (1994b). The tritone paradox: Some further geographical correlates. *Music Perception*, *12*, 125-136.
- DEUTSCH, D. (1996). Mothers and their children hear a musical illusion in strikingly similar ways. *Journal of the Acoustical Society of America*, *99*, 2482 [Abstract].
- 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.

- GLASS, I., & WOLLBERG, Z. (1983). Response of cells in the auditory cortex of awake squirrel monkeys to normal and reversed species-specific vocalizations. *Hearing Research*, **9**, 27-33.
- GRANTHAM, D. W., & WIGHTMAN, F. L. (1979). Auditory motion after-effects. *Perception & Psychophysics*, **26**, 403-408.
- GULICK, W. L., GESCHIEDER, G. A., & FRISINA, R. D. (1989). *Hearing: Physiological acoustics, neural coding, and psychoacoustics*. Oxford: Oxford University Press.
- HALL, J. W., & SODERQUIST, D. R. (1982). Transient complex and pure tone pitch changes by adaptation. *Journal of the Acoustical Society of America*, **71**, 665-670.
- HANLEY, T., SNIDECOR, J., & RINGEL, R. (1966). Some acoustic differences among languages. *Phonetica*, **14**, 97-107.
- MACKINNON, K. A. (1993). *The tritone paradox: Incidence in a student population, effects of music training, and perception versus vocal production response modes*. Unpublished BSc thesis, Acadia University.
- RAGOZZINE, F., & DEUTSCH, D. (1994). A regional difference in perception of the tritone paradox within the United States. *Music Perception*, **12**, 213-225.
- REPP, B. H. (1994). The tritone paradox and the pitch range of the speaking voice: A dubious connection. *Music Perception*, **12**, 227-255.
- REPP, B. H. (1997). Spectral envelope and context effects in the tritone paradox. *Perception*, **26**, 645-665.
- SHEPARD, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, **36**, 2346-2353.
- SHU, Z. J., SWINDALE, N. V., & CYNADER, M. S. (1993). Spectral motion produces an auditory after-effect. *Nature*, **364**, 721-723.
- WEINBERGER, N. M., & MCKENNA, T. M. (1988). Sensitivity of single neurons in auditory cortex to contour: Toward a neurophysiology of music perception. *Music Perception*, **5**, 355-390.
- WHITFIELD, I. C., & EVANS, E. F. (1965). Response of auditory cortical neurons to stimuli of changing frequency. *Journal of Neurophysiology*, **28**, 655-672.
- WILLEMS, N., COLLIER, R., & 'T HART, J. (1988). A synthesis scheme for British English intonation. *Journal of the Acoustical Society of America*, **84**, 1250-1261.
- YOST, W. A. (1994). *Fundamentals of hearing: An introduction* (3rd ed.). San Diego: Academic Press.

(Manuscript received June 28, 1996;
revision accepted for publication January 16, 1997.)