Music Perception Spring 1986, Vol. 3, No. 3, 275—280

A Musical Paradox

DIANA DEUTSCH
University of California, San Diego

A pattern of tones is explored that possesses some remarkable properties. It is heard as ascending when played in one key, yet as descending when played in another key. When the pattern is recorded on tape and then played back at different speeds, it is heard either as ascending or as descending depending on the speed of playback. As a further paradox, the pattern in any given key is heard as ascending by some listeners but as descending by others. Theoretical implications of these findings are discussed.

In common experience, when a melody is played in one key, and is then transposed to a different key, the perceived relationships between the tones are unchanged. In this respect, melodies are like visual shapes, which retain their perceptual identities when they are translated to different regions of the visual field (Von Ehrenfels, 1890). Indeed, the notion that a melodic pattern might be perceived as radically different under transposition appears as paradoxical as the notion that a visual shape might undergo a metamorphosis through being shifted to a different location in space.

The present study concerns a pattern with some exceptional properties. When played in one key it is heard as ascending, yet when played in a different key it is heard as descending instead. When a tape recording is made of this pattern, and it is played back at different speeds, the pattern is heard either as ascending or as descending depending on the speed of playback. To add to the paradox, the pattern in any given key is heard as ascending by some listeners, but as descending by others.

The pattern consists of two successively presented Shepard tones that are related by a half octave. As shown in the example in Figure 1, each tone consists of a set of octave-related sinusoids, the amplitudes of which are scaled by a fixed, bell-shaped spectral envelope. Such tones are well defined in terms of pitch class (C, C#, D, etc.), but are poorly defined in terms of height, since the usual cues for height attribution are missing.

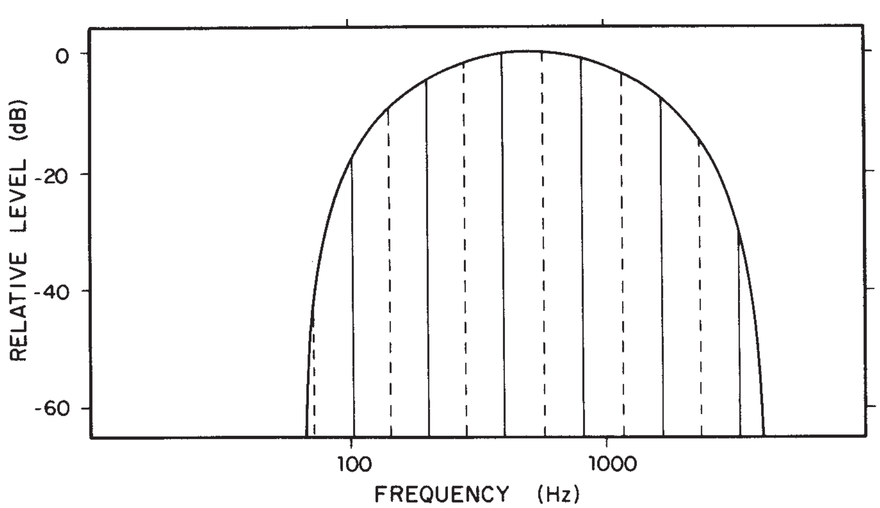

Fig. 1. Spectral representation of the tones used to generate one example of the paradoxical pattern. In this example, the spectral envelope is centered on C_5 (523 Hz). Dashed lines represent a tone of pitch class D, and solid lines represent a tone of pitch class G#. The two spectra are superimposed in the illustration, but the tones are played in succession.

Shepard (1964) has shown that when two such tones are played in succession, one hears either an ascending pattern or a descending one, depending on which is the shorter distance between the tones along the pitch class circle (Figure 2). So, for example, when D is played followed by D#, one hears the pattern as ascending, since the shorter distance here is clockwise. Similarly, when A is played followed by G, one hears the pattern as descending, since the shorter distance here is counterclockwise. These findings show that the perceptual system will, where appropriate, invoke proximity in making judgments of relative height for such tones. (See also Burns, 1981; Pollack, 1978; Risset, 1971; Schroeder, in press.)

But we can then ask, what happens when a pair of these tones is presented that are separated by exactly a half octave (or tritone), so that the same distance between them is traversed in either direction? What happens, for example, when C is played followed by F#, or A is played followed by D#, and so on? Will judgments of height here be ambiguous, or will some other principle be invoked to resolve the ambiguity? In Shepard's experiment, data were averaged over subjects and also over pitch classes, and for such data, ascending and descending judgments occurred equally often. This finding is consistent with a perceived ambiguity of height in this condition.

Let us, however, consider more closely the possibility that the perceptual system will not settle for ambiguity, but will instead invoke some other principle in making judgments of relative height. What cues would be available for it to draw on? As one possibility, reference could be made to the absolute positions of the tones along the pitch class circle. If tones in one region on the circle were tagged as higher, and tones in the opposite region as lower, the ambiguity would be resolved.

The present study was undertaken to examine this possibility. Subjects were presented with such tritone pairs, and they judged for each pair whether it formed an ascending or a descending pattern. The results were analyzed separately for each subject as a function of the pitch class of the first tone of the pair. As will be shown, the hypothesis was strikingly confirmed. For most subjects, judgments depended in a systematic fashion on the pitch class of the first tone, showing that tones in one region of the circle were heard as higher, and tones in the opposite region as lower. Yet the way in which the pitch-class circle was oriented with respect to height differed radically from one subject to another.

Method

Procedure

Subjects were tested in soundproof booths. On each trial a sequential pair of tones was presented, and the subjects judged whether it formed an ascending or a descending pattern. All tones were 500 msec in duration, and the second tone of each pair followed the first without pause. Each subject was tested in two sessions, which were held on different days, and the results from the two sessions were averaged. Each session consisted of 12 blocks of 12 trials each. Trials within blocks were separated by 5-sec intertrial intervals, and blocks were separated by 1-min pauses. There was a 5-min break between the sixth and seventh blocks. A few practice trials were given at the beginning of each session.

Equipment

Tones were generated on a VAX 11/780 computer, with the sound synthesis system developed by Moore (1982). They were recorded and played back on a Sony PCM-F1 digital audio processor, the output of which was passed through a Crown amplifier, and presented to subjects binaurally through Grason-Stadler TDH-49 headphones, at a level of approximately 72 dB SPL.

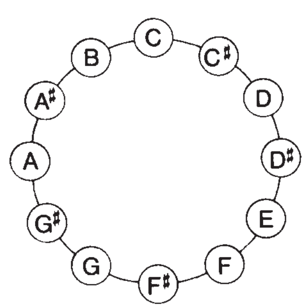

Fig. 2. The pitch-class circle.

Stimulus Parameters

Each tone consisted of six sinusoids that were separated by octaves, and the amplitudes of the sinusoids were determined by a fixed bell-shaped spectral envelope (Figure 1). A detailed specification of the shape of the envelope is given in Deutsch, Moore, and Dolson (1984). In order to control for possible effects based on the relative amplitudes of the sinusoids, all tone pairs were generated with envelopes placed at six different positions along the spectrum. The envelope peaks were spaced at half-octave intervals, specifically at $F\#_3$ (185 Hz), C_4 (262 Hz), $F\#_4$ (370 Hz), C_5 (523 Hz), $F\#_5$ (740 Hz), and C_6 (1047 Hz)[1]. Thus, for any given pitch class, the relative amplitudes of the sinusoidal components of tones generated under envelopes centered on C_4, C_5, and C_6 were identical to those for the pitch class a half-octave removed for tones generated under envelopes centered on $F\#_3$, $F\#_4$, and $F\#_5$. In addition, the use of envelopes with positions varying over a 2 1/2 octave range served to establish the generality of the findings for tone pairs that differed substantially in overall height.

Twelve tone pairs were generated under each of the six spectral envelopes, corresponding to the pitch-class pairings C–F#, C#–G, D–G#, D#–A, E–A#, F–B, F#–C, G–C#, G#–D, A–D#, A#–E, and B–F. This produced 72 tone pairs in all. The pairs were presented in blocks of 12, each block consisting of tones generated under one of the six envelopes, and containing one example of each of the pitch-class pairings. Within blocks, the 12 tone pairs were presented in either of two orders. These were essentially random, except that the same pitch classes did not occur in any two consecutive pairs. Twelve blocks were created altogether, with the two within-block orders employed once for each of the envelopes. The twelve blocks were presented in random order.

Subjects

Eight musically trained subjects with normal hearing participated in the experiment. Seven were students at the University of California, San Diego, and were paid for their participation. The author also served as subject.

Results and Discussion

The judgments of six of the subjects depended in a systematic fashion on the positions of the tones along the pitch-class circle. Those of the remaining two subjects did not show such a systematic relationship and are not further considered here.

Figure 3 displays, for each of the six subjects, the percentages of judgments that a tone pair formed a descending pattern, as a function of the pitch class of the first tone of the pair. From these judgments we can infer which tones were heard as higher and which as lower. It can be seen that striking differences between subjects emerged in the way that the pitch-class circle was oriented with respect to height. Subject FT, for example, consistently heard pitch classes B to E as higher and F# to A as lower. DD, in contrast, produced a curve that was almost the mirror image of that of FT, consistently hearing pitch classes F# to B as higher and C# to E as lower. TK showed an orientation with tones A to C as higher and C# to G as lower. MB showed a clear peak at B and a corresponding trough at F. For subject AS, A# to D# were higher, and E to A were lower. For subject MJD, F# to A were higher, and C to D# were lower.

It appears from these results that the orientation of the pitch-class circle varies in a haphazard fashion between listeners. This finding has the intriguing consequence that patterns composed from such tone complexes should be perceived by an audience in a large number of different ways. For example, the series of tone pairs E–A#, F#–C, D#–A, G–C#, would for the case of listeners such as FT be perceived as in Figure 4a. However, for the case of listeners such as DD, the identical series of tone pairs would be perceived as in Figure 4b instead!

A particularly striking demonstration of this paradox may be produced by tape recording a number of exemplars of the same tone pair and then playing the tape back at different speeds. In particular, if the tape is first played at normal speed, and is then sped up so that the tones are transposed up a half octave, listeners who first heard the pattern as ascending will now hear it as descending, and listeners who first heard the pattern as descending will now hear it as ascending!

A related association between pitch class and perceived height has recently been obtained by Deutsch et al. (1984) employing two-part patterns consisting of Shepard tones. Such patterns were found to be perceived quite differently depending on which key they were in, so that transposition led to an interchange of voices. The dependence on key did not shift with shifts in the position of the spectral envelope, showing that the effect was not a result of judgments based on the individual spectral components.

1. C_4 corresponds to Middle C, C_5 to the octave above Middle C, and so on.

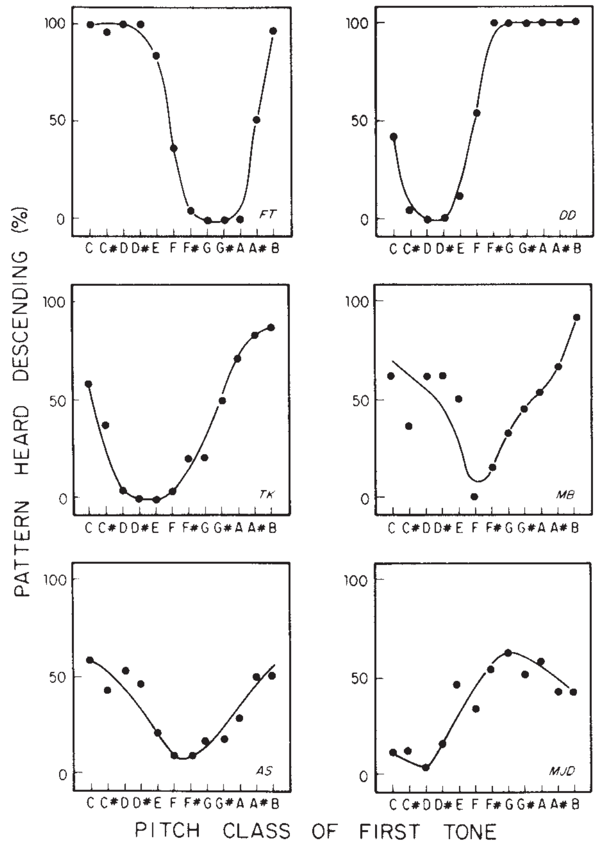

Fig. 3. Percentages of judgments that a tone pair formed a descending pattern, plotted as a function of the pitch class of the first tone of the pair. Data from individual subjects are displayed, averaged over two sessions.

Fig. 4. Representation of the identical series of tone pairs, as perceived by two different subjects. In general, such a series, when constructed as described here, will be perceived by an audience in a large number of different ways.

The finding that listeners are able to refer to the absolute positions of tones along the pitch-class circle in making judgments of height shows that the possession of absolute pitch is, in a sense, considerably more common than has been assumed (a point made recently in a different context by Terhardt & Ward, 1982 and Terhardt & Seewann, 1983).[2] The extent to which this principle is invoked in determining the perceived heights of naturally occurring sounds remains to be investigated.[3]

2. Only one subject in the present study possessed absolute pitch, as traditionally determined by the ability to attach verbal labels to notes played in isolation.

3. This study was reported in an Invited Address to the Ninety-Third Annual Convention of the American Psychological Association, Los Angeles, August, 1985. The work was supported in part by USPHS Grant MH-21001. I am particularly grateful to F. Richard Moore for the use of the Computer Audio Research Laboratory at the UCSD Center for Music Experiment, to F. Richard Moore and Mark Dolson for software used in generating the sound patterns, and to Lee Ray for technical assistance.

References

Burns, E. Circularity in relative pitch judgements for inharmonic tones: The Shepard demonstration revisited, again. *Perception and Psychophysics*, 1981, **30**, 467-472.

Deutsch, D., Moore, F. R., & Dolson, M. Pitch classes differ with respect to height. *Music Perception*, 1984, **2,** 265-271.

Moore, F. R. The computer audio research laboratory at UCSD. *Computer Music Journal*, 1982, **6**, 18-29.

Pollack, I. Decoupling of auditory pitch and stimulus frequency: The Shepard demonstration revisited. *Journal of the Acoustical Society of America*, 1978, **63**, 202-206.

Risset, J. C. Paradoxes de hauteur: Le concept de hauteur sonore n'est pas le meme pour tout le monde. *Seventh International Congress of Acoustics*, Budapest, 1971, p. 20, S10.

Schroeder, M. R. Auditory paradox based on fractal waveform. *Journal of the Acoustical Society of America*, in press.

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

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

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

Von Ehrenfels, C. Über Gestaltqualitaten. *Vierteljahrschrift für Wissenschaftliche Philosophie*, 1890, **14**, 249-292.