Music Perception
Winter 1984, Vol. 2, No. 2, 265–271

Pitch Classes Differ with Respect to Height

DIANA DEUTSCH, F. RICHARD MOORE,
& MARK DOLSON
University of California, San Diego

The pitch of a tone is considered to consist of two components: the rectilinear component of height and the circular component of chroma, or pitch class. In an experiment employing simultaneous sequences of Shepard tones, it is shown that tones in different positions on the chroma circle differ in height when generated under the same spectral amplitude envelope. Further, the directions of these differences remain constant when the envelope is centered at different positions along the spectrum, resulting in clear differences in the overall heights of the patterns.

THE pitch of a tone may be described as consisting of two components. First, the rectilinear component of *height* corresponds to the overall position of the tone on a continuum from high to low. Second, the circular component of *chroma* (Bachem, 1948; Shepard, 1964) or *pitch class* (Babbitt, 1960) corresponds to the position of the tone within the octave. Note names in the Western musical scale repeat at octave intervals, reflecting the circularity of the chroma dimension. Thus the note names C_3, C_4, and C_5 correspond to the tones of the same chroma that are placed in three different octaves. The note names $D\sharp_4$, E_4, and G_4 correspond to tones of different chroma that are placed in the same octave. The dimensions of height and chroma are considered to be orthogonal. Thus the statement "B is higher than D" is considered meaningful only when octaves are also specified: B_4 is higher than D_4 but lower than D_5.[1]

With sounds produced by natural instruments, as with most electronically produced sounds, shifts in chroma are always correlated with differences in spectrum, resulting in differences in perceived height. However, Shepard (1964) proposed a method of physically dissociating these two dimensions. He generated a series of tones, each of which consisted of a set

Requests for reprints may be sent to Diana Deutsch, Department of Psychology, C-009, University of California at San Diego, La Jolla, California 92093. Cassettes of sound examples are available at cost by writing to the first author.

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

of sinusoids that were separated by octaves, with their amplitudes determined by a fixed bell-shaped spectral envelope. When listeners were presented with ordered pairs of such tones, and were asked to judge which member of each pair was higher in pitch, the averaged data indicated that such judgments were circular. Thus, for example, C would be judged as lower than D♯, D♯ as lower than F♯, F♯ as lower than A, and A as lower than C. It was concluded that, for such tones, the dimension of height was entirely suppressed, with chroma alone remaining.

In Shepard's experiment, however, judgments were overwhelmingly influenced by the perceptual tendency to structure pitch sequences in accordance with proximity (see also Deutsch, 1975, 1982; Bregman, 1978; Bregman & Campbell, 1971; Van Noorden, 1975). Thus, for example, the pattern (C–C♯) would always be perceived as ascending by one semitone, never as descending by 11 semitones. Similarly, the pattern (G–F) would always be perceived as descending by two semitones, never as ascending by 10 semitones. The question then remains as to whether such tones might be found to differ in height if this factor were controlled for. We here present findings showing that, at least for two-part patterns, tones in different positions on the chroma circle do indeed differ in height when generated under the same spectral amplitude envelope. Further, the directions of these differences remain constant when the envelope is centered at different positions along the spectrum, resulting in clear differences in the overall height of the patterns.

The experiment employed two-part patterns composed of Shepard tones. The first was in the key of C major, and consisted of the sequence (D–E–F) presented simultaneously with the sequence (B–A–G). The second was an exact transposition of the first to the key of F♯ major, and so consisted of the sequence (G♯–A♯–B) presented simultaneously with the sequence (E♯–D♯–C♯). Previous work (Deutsch & Moore, 1984) has shown that when such a pattern is perceived unambiguously, it is perceptually organized in two parts in accordance with pitch proximity. Thus the listener perceives one melodic line that ascends by a minor third, together with another that descends by a major third. However, the higher line may be heard as ascending and the lower as descending, or alternatively the higher line may be heard as descending and the lower as ascending. These two organizations are illustrated in Figure 1.[2]

On each trial one of the above patterns was presented, and the listener judged whether the higher line ascended or descended. From these judgments it was inferred which tones were heard as higher and which as lower.

2. In the experiment of Deutsch and Moore (1984), subjects reported the patterns that they heard in musical notation. The findings of the present experiment were also confirmed when subjects were asked to notate the patterns.

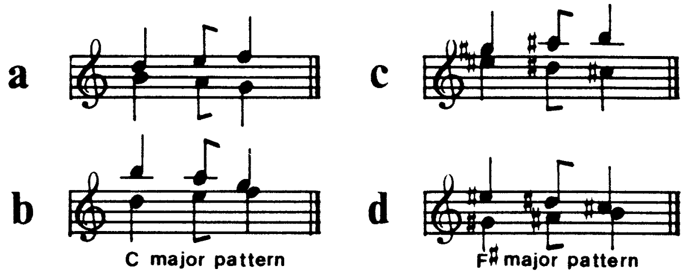

Fig. 1. Pitch patterns employed in the experiment, each notated in accordance with two different perceptual organizations. (a,c) Higher line perceived as ascending and lower line as descending. (b,d) Higher line perceived as descending and lower line as ascending.

Thus if the C major pattern was heard with the higher line ascending, this meant that tones D, E, and F were heard as higher and tones B, A, and G as lower. Conversely, if this pattern was heard with the higher line descending, this meant that tones B, A, and G were heard as higher, and tones D, E, and F as lower. Similarly, if the F♯ major pattern was heard with the higher line ascending, this meant that tones G♯, A♯, and B were heard as higher, and tones F (i.e., E♯), D♯, and C♯ as lower. Conversely, if this pattern was heard with the higher line descending, this meant that tones F, D♯, and C♯ were heard as higher and the tones G♯, A♯, and B as lower.

In order to ensure that the obtained effects were based on pitch class and not simply on pitch, the two patterns were each generated under four different spectral amplitude envelopes. These were centered on C_4, $F♯_4$, C_5, and $F♯_5$. Tones generated under these different envelopes sound clearly different in height. Thus tones with envelopes centered on C_4 and C_5 appear as in different octaves, as do tones with envelopes centered on $F♯_4$ and $F♯_5$. Further, any differences in the relative amplitudes of the components of tones generated under spectra centered on C_4 and C_5 were counterbalanced by differences in the opposite direction for tones generated under spectra centered on $F♯_4$ and $F♯_5$.

Method

Stimulus parameters

Tones were generated on line by a VAX 11/780 computer, with the music sound synthesis system developed by one of the authors (Moore, 1982). For each pattern, each tone consisted of the sum of a collection of octave-related sinusoidal components, with the amplitude of each component scaled by a fixed spectral amplitude envelope. The detailed shape and bandwidth of the spectral envelope determined the amplitude of each of the sinusoidal components according to their individual frequencies.

For this experiment, the spectral amplitude envelope consisted of a "raised, inverted" cosine curve logarithmically scaled to lie between f_{min} and f_{max} according to

$$A(f) = 0.5 - 0.5\cos(\frac{2\pi}{\gamma} \log_\beta [1 + h(f)(\beta^\gamma - 1)])$$

where:

$A(f)$ is the relative amplitude of a sinusoid at frequency f Hz,
β is the frequency ratio between adjacent sinusoidal components (for example, $\beta = 2$ for octave spacing),
γ is the number of β-cycles between f_{min} and f_{max},
$h(f)$ is an "interpolating" function that goes linearly from 0 to 1 as f goes from f_{min} to f_{max}, that is,

$$h(f) = \frac{f - f_{min}}{f_{max} - f_{min}}$$

We chose f_{max} to be at 6 octaves (or more generally, β-cycles) above f_{min}. Under these conditions, at most 7 octave-related sinusoids would "fit" under the spectral envelope curve. In terms of musical pitch, if f_{min} corresponds to c_0 and f_{max} corresponds to c_6 (six octaves higher), then the Shepard tone with pitch class c would contain sinusoidal components at frequencies corresponding to $c_0, c_1, \ldots, c_6$. Note that the lowest and highest component amplitudes will be zero in this case. To generate the Shepard tone with pitch class d, only the frequencies corresponding to pitches d_0 through d_5 "fit" under the same spectral amplitude envelope, all with nonzero amplitude. The number of components with nonzero amplitude was therefore equal to either $\gamma - 1$ or $\gamma - 2$, depending on the pitch class of the generated Shepard tone (Figure 2).

The two pitch patterns shown on Figure 1 were each generated under spectral amplitude envelopes centered on C_4, $F\sharp_4$, C_5, and $F\sharp_5$. For these, f_{min} was 32.7, 46.2, 65.4, and 92.5 Hz, respectively. Patterns were presented to listeners binaurally through earphones (Grason-Stadler TDH 49) at an amplitude of approximately 72 dB SPL.

For each pattern, the durations of the three successive tones were 500, 250, and 500 msec. Each pattern was presented three times on each trial, with presentations separated by 750-msec pauses. Trials were separated by 10-second intertrial intervals, during which the listeners made their judgments.

Procedure

Subjects were individually tested, each in two sessions. In the first session, the C major and F♯ major pitch patterns were presented, each generated under the C_5 and $F\sharp_5$ spectral envelopes. The four stimulus patterns were delivered four times in succession, each time in a different random order. In the second session, the C major and F♯ major pitch patterns were again presented, now generated under the C_4 and $F\sharp_4$ spectral envelopes. These four stimulus patterns were also delivered four times in succession, each time in a different random order. The subject judged on each trial whether the higher line ascended or descended.

Subjects

The study of Deutsch and Moore (1984) had shown that there were consistent differences between listeners in terms of how such patterns are perceived. For this reason, three groups of subjects were selected on the basis of their percept of the C major pattern generated under the spectral amplitude envelope centered on C_5. Three were selected who consistently heard this pattern as with the higher line ascending, so that their percepts were as in Figure 1a (Type A). Three were selected who heard this pattern as with the higher line descending, so that their percepts were as in Figure 1b (Type B). Three were selected who heard the pattern ambiguously (Type C). All subjects had received musical training.

DEUTSCH, DIANA, ***Pitch Classes Differ with Respect to Height***, Music Perception, 2:2 (1984:Winter) p.265

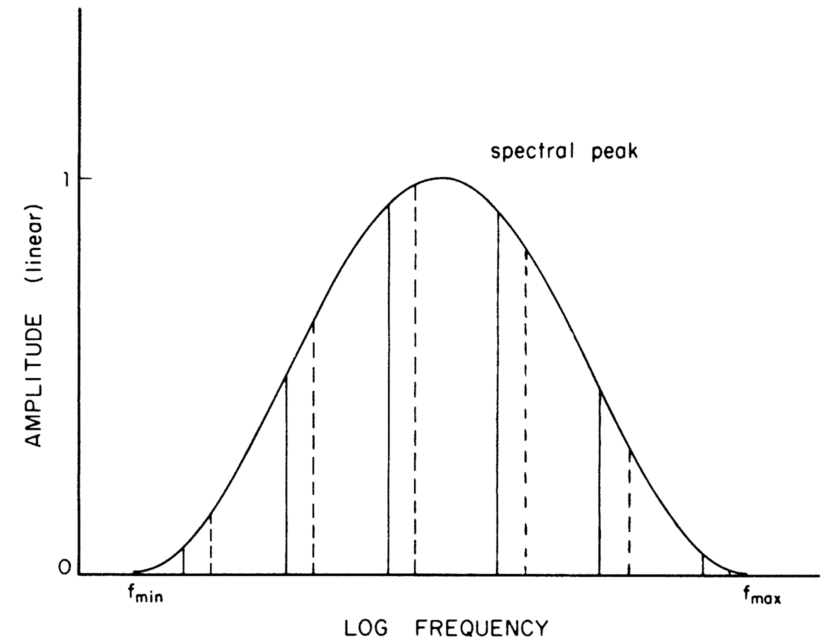

Fig. 2. Representation of Shepard tones used in the experiment. Each tone consisted of a set of sinusoids that were separated by octaves, with their amplitudes determined by a fixed spectral amplitude envelope. Solid lines produced a tone of pitch class A, and dashed lines produced a tone of pitch class B. Tones within any given pattern were generated under the same spectral amplitude envelope. However, for different patterns the envelope was rigidly shifted in log frequency, so that the spectral peak was at C_4, $F\sharp_4$, C_5, or $F\sharp_5$.

Results and Discussion

The results of the experiment are shown in Table 1. It can be seen that pronounced effects of position on the chroma circle were obtained.

Type A subjects reported the C major pattern as with the higher line ascending, so that they heard it with pitch classes D, E, and F as higher and pitch classes B, A, and G as lower (i.e., as in Figure 1a). Further, they reported the F♯ major pattern as with the higher line descending, so that they heard it with pitch classes F (i.e., E♯), D♯, and C♯ as higher and pitch classes G♯, A♯, and B as lower (i.e., as in Figure 1d). These percepts were obtained with patterns generated under all four spectral amplitude envelopes.

Type B subjects obtained percepts that were the converse of Type A. They reported the C major pattern as with the higher line descending, so that they heard it with pitch classes B, A, and G as higher and pitch classes D, E, and F as lower (i.e., as in Figure 1b). Further, they reported the F♯ major pattern as with the higher melodic line ascending, so that they heard it with pitch classes G♯, A♯ and B as higher and F, D♯, and C♯ as lower (i.e., as in

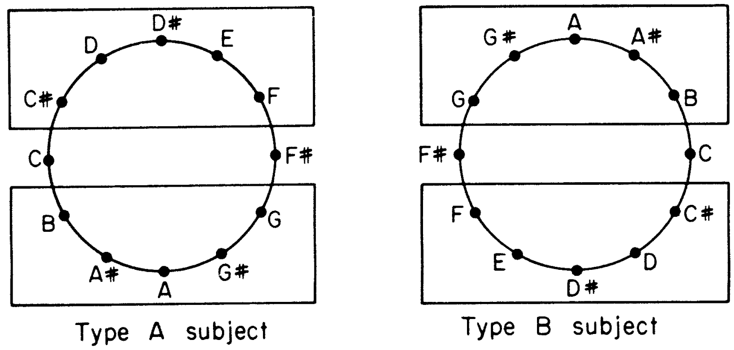

Fig. 3. "Upright" orientations of the chroma circle, for Type A and Type B subjects. Tones bounded by upper rectangle were heard as higher, and tones bounded by lower rectangle as lower.

TABLE 1
Percentages of Reports of Each Stimulus Pattern as with the Higher Line Ascending

Type A Subjects

	Spectral Peak C_4	$F\#_4$	C_5	$F\#_5$
C major	100	100	92	92
F# major	0	0	0	0

Type B Subjects

	Spectral Peak C_4	$F\#_4$	C_5	$F\#_5$
C major	0	0	0	0
F# major	100	100	83	92

Type C Subjects

	Spectral Peak C_4	$F\#_4$	C_5	$F\#_5$
C major	25	50	42	58
F# major	33	42	50	50

NOTE: C major pattern: ascending line (D–E–F), descending line (B–A–G). F# major pattern: ascending line (G#–A#–B), descending line (E#–D#–C#).

Figure 1c). These percepts were also obtained for patterns generated under all four envelopes.

The percepts of Type A and Type B subjects are illustrated in Figure 3. The chroma circle is "upright" in one orientation for Type A subjects, and in the opposite orientation for Type B subjects. For both groups, this effect was so strong as to override considerations based on tonality. Although the C major and F♯ major patterns are exact transpositions of each other, these subjects never perceived them as such, since the difference in key was associated with an interchange of voices.

Type C subjects heard both patterns ambiguously, and this effect was also consistent across differences in spectra.

The findings reported here are at the phenomenological level very compelling. They provide a paradox for theories of pitch perception, which so far have focused on the height dimension, and have not considered the possibility of effects based primarily on pitch class (Goldstein, 1973; Terhardt, Stoll, & Seewann, 1982; Wightman, 1973). The extent to which phenomena such as those described here occur in perception of other complex sounds remains to be investigated.[3]

References

Babbitt, M. Twelve-tone invariants as compositional determinants. *The Musical Quarterly*, 1960, ***46***, 246–259.

Bachem, A. Note on Neu's review of the literature on absolute pitch. *Psychological Bulletin*, 1948, ***45***, 161–162.

Bregman, A. S. The formation of auditory streams. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, New Jersey: Erlbaum, 1978.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and the perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, ***89***, 244–249.

Deutsch, D. Musical illusions. *Scientific American*, 1975, ***233***, 92–104.

Deutsch, D. Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press, 1982.

Deutsch, D., & Moore, F. R. A spectral contrast illusion. *Journal of the Acoustical Society of America*, 1984, ***176***, S77.

Goldstein, J. L. An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America*, 1973, ***54***, 1496–1516.

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

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

Terhardt, E., Stoll, G., & Seewann, M. Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions. *Journal of the Acoustical Society of America*, 1982, ***71***, 671–678.

Van Noorden, L. P. A. S. *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation, Technische Hogeschoel Eindhoven, The Netherlands, 1975.

Wightman, F. L. The pattern-transformation model of pitch. *Journal of the Acoustical Society of America*, 1973, ***54***, 407–416.

3. This work was supported by USPHS Grant MH-21001 to the first author.