Diana Deutsch

Pitch Class and Perceived Height: Some Paradoxes and Their Implications

Much lively debate exists concerning the framework within which musical materials should be considered as organized, classified, and abstracted. A number of concepts are, however, taken as axiomatic since they appear on common-sense grounds to be beyond dispute. One such is the concept of a musical note as having a pitch class and a register appropriate to its name (e.g., C_4, D_5, and so on). Another is the concept of invariance under transposition: Although one might argue about details, it is taken as generally self-evident that a musical passage retains its identity when played in different keys.

These two concepts are related to yet another which is also generally taken as axiomatic, namely the orthogonality of pitch class and perceived height. If a musician were faced with the question "Which note is higher, C-sharp or G?," he would probably reply that the question was nonsensical: one would have to know; *which* C-sharp and *which* G before a meaningful answer could be given. This essay presents some surprising findings showing that, at least under certain conditions, pitch class and height are not orthogonal; rather the perceived height of a tone can be shown to be related in an orderly fashion to its position along the pitch class circle. The theoretical and practical implications of these findings are discussed.

In E. Narmour and R. Solie (Eds) *Explorations in Music, the Arts and Ideas; Essays in Honor of Leonard B. Meyer* Pendragon Press, Stuyvesant, 1988.

PITCH AS A GEOMETRICALLY REGULAR HELIX

The view that pitch class and height are orthogonal forms the basis of the model illustrated in figure 1. Pitch is here assumed to vary along both a circular dimension of pitch class and also a monotonic dimension of height. It can thus be represented as a geometrically regular helix, in which the entire structure maps into itself by transposition.[1] One can readily see that in this representation, any musical interval is represented by pairs of points that are separated by the same distance from each other. Further, the strong perceptual similarity between tones that are related by octaves is captured by their being depicted in relatively close spatial proximity.

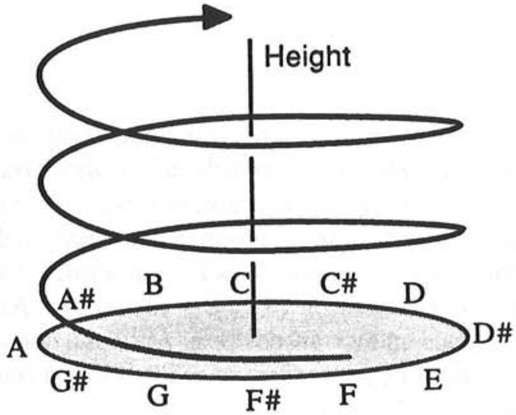

Figure 1. Pitch as a geometrically regular helix. (Adapted from Shepard, 1965.)

This helical model, as pointed out by Shepard,[2] leads to the intriguing possibility that, by suppressing the monotonic component of height, leaving only the circular component of pitch class, all tones that are an octave apart could be mapped onto the same tone, which would then

[1]See, for example, Roger N. Shepard, "Circularity in judgments of relative pitch," *Journal of the Acoustical Society of America*, 36 (1964): 2346-2353; "Approximation to Uniform Gradients of Generalization by Monotone Transformations of Scale," in *Stimulus Generalization*, ed. D. I. Mostofsky (Stanford, 1965): 94-110; and "Structural Representations of Musical Pitch," in *The Psychology of Music*, ed. Diana Deutsch (New York, 1982): 344-390; and also earlier work by M. W. Drobisch, "Über musikalische Tonbestimmung und Temperatur," *Abhandl. Math. Phys. Kl. Konigl. Sachs. Ges. Wiss*, 4 (1855): 1-120.
[2]Shepard, "Circularity."

have a clearly determined pitch class but an indeterminate height. The tonal helix would thus be collapsed into a circle, and judgments of pitch would be expected to be completely circular.

In order to test this prediction, Shepard performed an experiment in which he employed a specially contrived set of tones. Each tone consisted of ten sinusoidal components that were separated by octaves. The amplitudes of these components were scaled by a fixed, bell-shaped spectral envelope, so that those in the middle of the musical range were loudest, with the lowest and highest falling off below the threshold of audibility. Such tones are heard as well defined in terms of pitch class but somewhat ambiguously in terms of height.

Listeners were presented with ordered pairs of such tones and were asked to judge whether they formed ascending or descending series. When the tones within a pair were separated by a small distance along the pitch class circle, these judgments of relative height were found to be entirely dependent on proximity.[3] Thus, for example, the pattern C – C-sharp would always be heard as ascending, as would the patterns D – D-sharp, F – F-sharp, and B – C. Analogously, the pattern C-sharp – C would always be heard as descending; and so on.[4] When the tones within a pair were separated by a larger distance along the pitch-class circle, the tendency to follow by proximity gradually decreased. When the tones within a pair were separated by exactly a half-octave, ascending and descending judgments occurred equally often.

Shepard concluded that for such octave-related complexes, the monotonic dimension of height was indeed suppressed, leaving only

[3]Proximity is one of the principles of perceptual organization described by the Gestalt psychologists at the turn of the century. These principles have been shown to be prominently involved in the organization of tonal music. See particularly Leonard B. Meyer, *Emotion and Meaning in Music* (Chicago, 1956); and *Explaining Music: Essays and Explorations* (Berkeley, 1973).

[4]For other experimental work showing that proximity is invoked in making judgments of relative height, see Jean-Claude Risset, "Paradoxes de hauteur: Le concept de hauteur sonore n'est pas le meme pour tout le monde," paper presented to the *Seventh International Congress of Acoustics*, (Budapest,1971); Irwin Pollack, "Decoupling of Auditory Pitch and Stimulus Frequency: The Shepard Demonstration Revisited," *Journal of the Acoustical Society of America*, 63 (1978): 202-206; Edward Burns, "Circularity in Relative Pitch Judgments for Inharmonic Tones: The Shepard Demonstration Revisited, Again," *Perception and Psychophysics* 30 (1981): 467-472; Manfred R. Schroeder, "Auditory Paradox Based on Fractal Waveform," *Journal of the Acoustical Society of America*, 79 (1986): 186-188; Ryunen Teranishi, "Endlessly rising or falling chordal tones which can be played on the piano; another variation of the Shepard demonstration," paper presented to the *12th International Congress of Acoustics*, Toronto, 1986; and Kazuo Ueda and Kengo Ohgushi, "Perceptual components of pitch: Spatial representation using a multidimensional scaling technique," *Journal of the Acoustical Society of America*, 82 (1987): 1193-1200.

the circular dimension of pitch class. However, there are problems with this interpretation. In the case where the tones within a pair were related by close proximity, other factors that might have given rise to differences in perceived height would have been masked. Indeed, it has been shown in other contexts that proximity can override cues which would otherwise be operating to organize pitch materials.[5] Further, Shepard obtained his results by averaging over pitch classes, so that any relationship between pitch class and height would have been lost in the averaging process. The issue of orthogonality was thus left unresolved in his study.

THE PRESENT EXPERIMENTS

The present experiments were undertaken to explore the relationship between pitch class and perceived height where proximity could not be used as a cue. They examined the question of whether judgments of height would here be completely ambiguous, as predicted from the helical model, or whether some other principle might be invoked to resolve the ambiguity. A number of different types of pattern were explored, and in all cases, striking and paradoxical findings were obtained. The perceived heights of tones in such patterns were found to vary in an orderly fashion depending on their positions along the pitch class circle. As a result, when the patterns were transposed, entirely different configurations were perceived. These findings therefore constituted violations of the principle of invariance under transposition. As a further unexpected finding, the form of relationship between pitch class and perceived height varied substantially from one listener to another, so that any given pattern was perceived by different listeners in radically different ways.

[5]The scale illusion provides a particularly strong example of the overriding influnce of pitch proximity in the organization of musical materials. Here two series of tones emanate simultaneously from different regions of space. These are perceptually reorganized so as to create the illusion that tones in one pitch range are emanating from one region, and tones in a different pitch range from the other region. See Diana Deutsch, "Two-channel listening to musical scales," *Journal of the Acoustical Society of America* 157 (1975): 1156-1160; "Musical Illusions,", *Scientific American*, 233 (1975): 92-104. The principle underlying the scale illusion has been found to operate for a variety of musical patterns, and under a number of listening conditions. See David Butler, "A Further Study of Melodic Channeling," *Perception and Psychophysics* 25 (1979): 264-268; "Melodic Channeling in Musical Environment," paper presented at the *Research Symposium on the Psychology and Acoustics of Music*, Kansas (1979); and Diana Deutsch, "Dichotic Listening to Melodic Patterns and Its Relationship to Hemispheric Specialization of Function," *Music Perception* 3 (1985): 127-154.

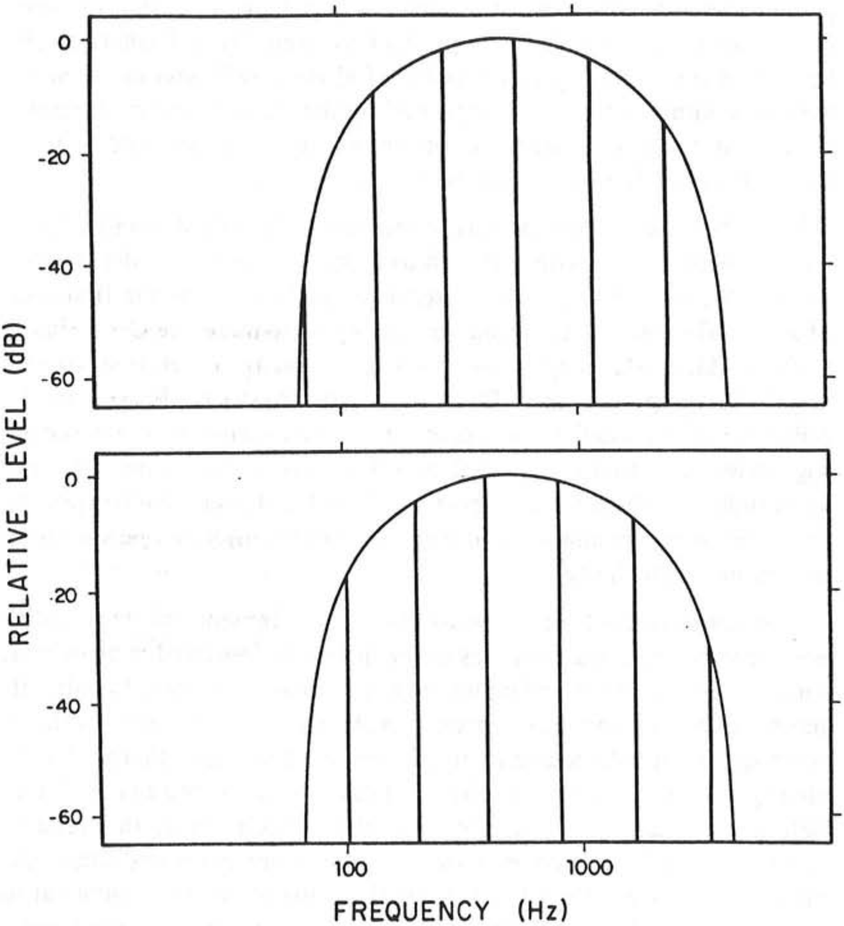

Figure 2. Representation of the spectral composition of a tone pair giving rise to the tritone paradox. In this case the spectral envelope is centered at C_5. Upper graph represents a tone of pitch class D, and lower graph a tone of pitch class G-sharp.

The Tritone Paradox

The first pattern to be described consisted of two successively presented tones which were related by a half-octave, and so were diametrically opposed along the pitch-class circle. Thus C-sharp might be presented followed by G, or A-sharp followed by E, and so on. Since the tones within a pair were separated by the same distance along the pitch-class circle in either direction, proximity could not here be used as a cue in making judgments of relative height.

In the first experiment on this phenomenon,[6] musically trained subjects were presented with just such tone pairs in random order, so that each of the twelve pitch classes served equally often as the first tone of a pair. The pitch class pairings employed were therefore C – F-sharp, C-sharp – G, D – G-sharp, D-sharp – A, E – A-sharp, F – B, F-sharp – C, G – C – G-sharp, G-sharp – D, A – D-sharp, A-sharp – E, and B – F. Subjects judged whether each pair formed an ascending or a descending series. All tones consisted of six octave-related sinusoids, the amplitudes of which were determined by a fixed, bell-shaped spectral envelope. A representation of the spectral composition of one such tone pair is given in figure 2.

In order to control for the possibility that judgments might be influenced by the relative amplitudes of the different sinusoidal components, tone pairs were generated under envelopes that were placed at six different positions along the spectrum. As shown in figure 3, the envelopes were spaced at half-octave intervals, centered at C_6, F-sharp$_5$, C_5, F-sharp$_4$, C_4, and F-sharp$_3$, so that the peaks varied over a two and one-half octave range.[7] Thus, for any given pitch class, the relative amplitudes of the components of tones that were generated under the envelopes shown on the left side of the illustration were identical to those for the pitch class a half octave removed that were generated under

[6]Diana Deutsch, "A musical paradox," *Music Perception* 3 (1986): 275-280.

[7]In this and the other experiments described in this chapter, tones were generated on a VAX 11/780 computer, interfaced with a DSC-200 Audio Data Conversion System, and using the music sound synthesis software (see F. Richard Moore, "The computer audio research laboratory at UCSD," *Computer Music Journal* 6 (1982): 18-29. The sounds were recorded and played back on a Sony PCM-F1 digital audio processor, the output of which was routed through a Crown amplifier and presented to subjects binaurally through headphones (Grason-Stadler TDH-49) at an approximate loudness level of 72 dB SPL. The convention in notation followed here is that C_4 corresponds to middle C, C_5 to the octave above middle C, and so on.

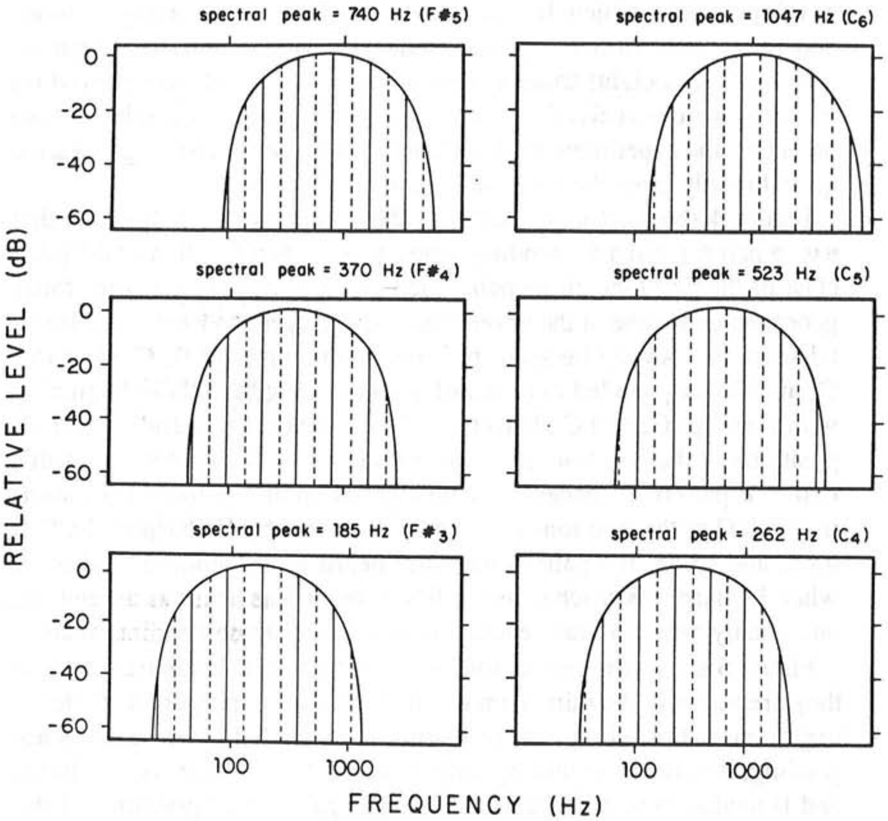

Figure 3. Representation of the spectral compositions of tones comprising the D – G-sharp pattern, generated under six spectral envelopes. Dashed lines indicate tones of pitch class D, and solid lines tones of pitch class G-sharp. The spectra of the two sets of tones are superimposed in the illustration, although the tones were presented in succession.

the envelopes shown on the right side. All twelve pitch-class pairings were generated under each of the six envelopes.[8]

At the phenomenological level, tones generated under the different envelopes sounded clearly diffferent in height. Roughly, the perceived height of one such tone corresponded to the center of the spectral envelope under which it was generated. Thus, for example, tones generated under an envelope centered on C_5 were perceived as approximating C_5 in height; tones generated under an envelope centered on F-$sharp_3$ were perceived as approximating F-$sharp_3$ in height; and so on. Thus, the experiment explored tones whose perceived heights varied quite broadly over the most salient musical range.

Figure 4 shows, for one subject, the percentage of judgments that a tone pair formed a descending series, plotted as a function of the pitch class of the first tone of the pair. Each graph plots judgments for tones generated under one of the envelopes, and averaged over two experimental sessions . As can be seen, patterns beginning with B, C, C-sharp, D, and D-sharp tended to be heard as descending, and those beginning with F-sharp, G, and G-sharp tended to be heard as ascending, for all positions of the spectral envelope. In other words, if we think of this two-tone pattern as successively transposed up in semitone steps, starting with C as the first tone of the pair, followed by C-sharp as the first tone, and so on, the pattern was first heard as descending, and then, when F-sharp was reached as the first tone, it was heard as ascending, and finally when B was reached it was heard as descending again.

Figure 5 shows the results for a different subject. It can be seen that they are virtually the mirror-image of those shown in figure 4. Patterns beginning with C-sharp, D, D-sharp, and E tended to be heard as ascending, and those beginning with F-sharp, G, G-sharp, A, A-sharp, and B tended to be heard as descending, again for all positions of the spectral envelope. So, thinking of the pattern as successively transposed up in semitone steps beginning with C-sharp as the first tone of a pair, the pattern was first heard as ascending, and then, when F-sharp was reached, it was heard as descending; and so on.

Figure 6 shows the results for these two subjects together, taking as an example judgments with the spectral envelope centered at F-$sharp_4$.l It can be seen that as the pattern was transposed up in semitone steps, it was first heard one way and then it was heard as inverted. But for

[8]All tones were 500 msec in duration, and there were no pauses between tones within a pair. Tone pairs were presented in blocks of twelve, with five-second pauses between pairs within a block, and one-minute pauses between blocks. Each block consisted of tones generated under one of the six envelopesand contained one example of each of the pitch-class pairings.

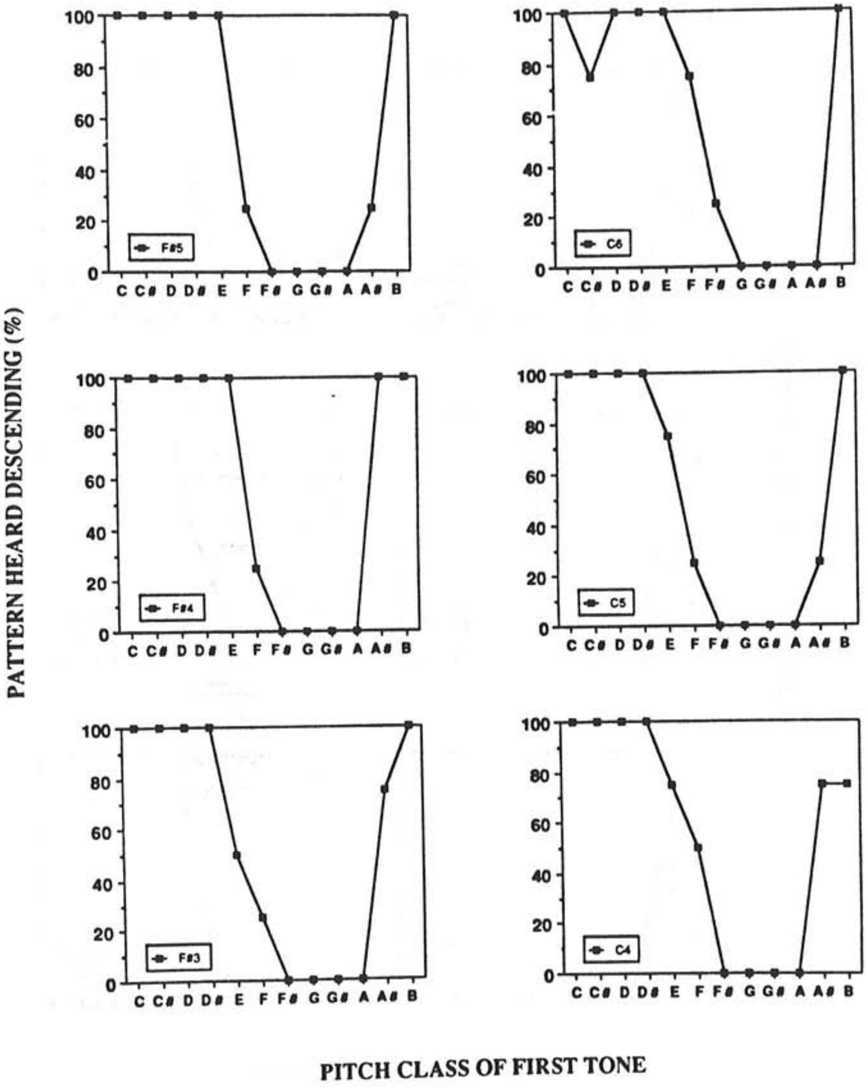

Figure 4. Percentages of judgments that a tone pair formed a descending series, plotted as a function of the pitch class of the first tone of the pair. Results from a first subject are here displayed, for tones generated under each of the six spectral envelopes shown in figure 3. Symbols in boxes indicate the peaks of the spectral envelopes. (Data from Deutsch, 1986.)

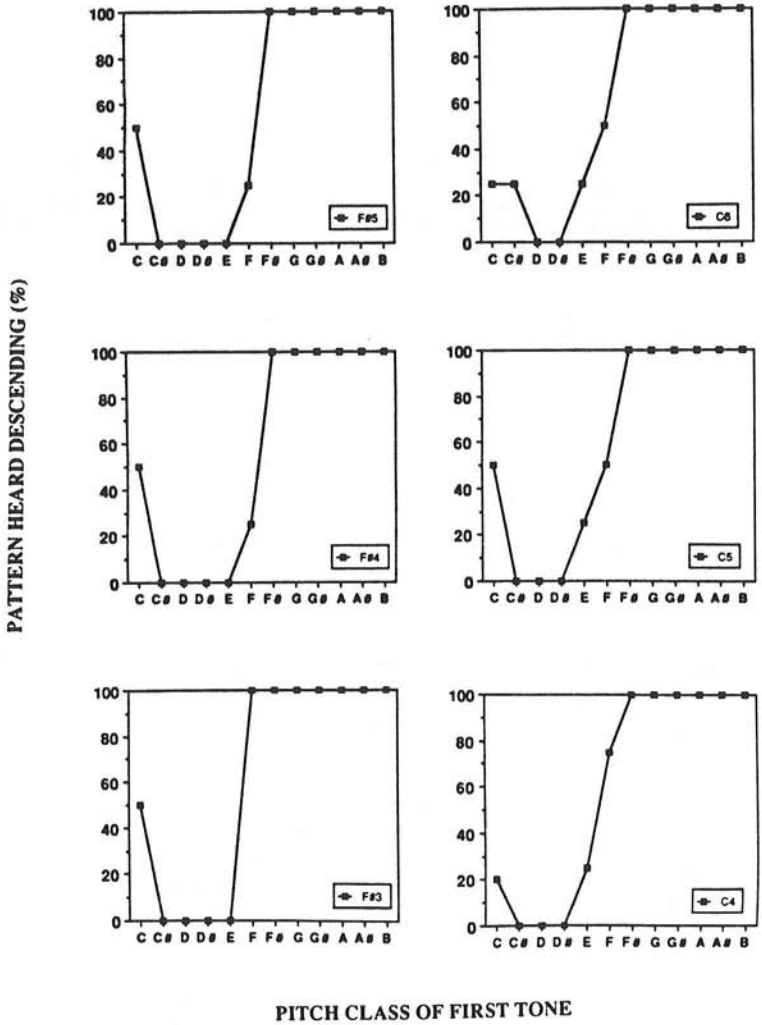

Figure 5. Percentages of judgments that a tone pair formed a descending series, plotted as a function of the pitch class of the first tone of the pair. Results from a second subject are here displayed, for tones generated under each of the six spectral envelopes shown in figure 3. Symbols in boxes indicate the peaks of the spectral envelopes. (Data from Deutsch, 1986.)

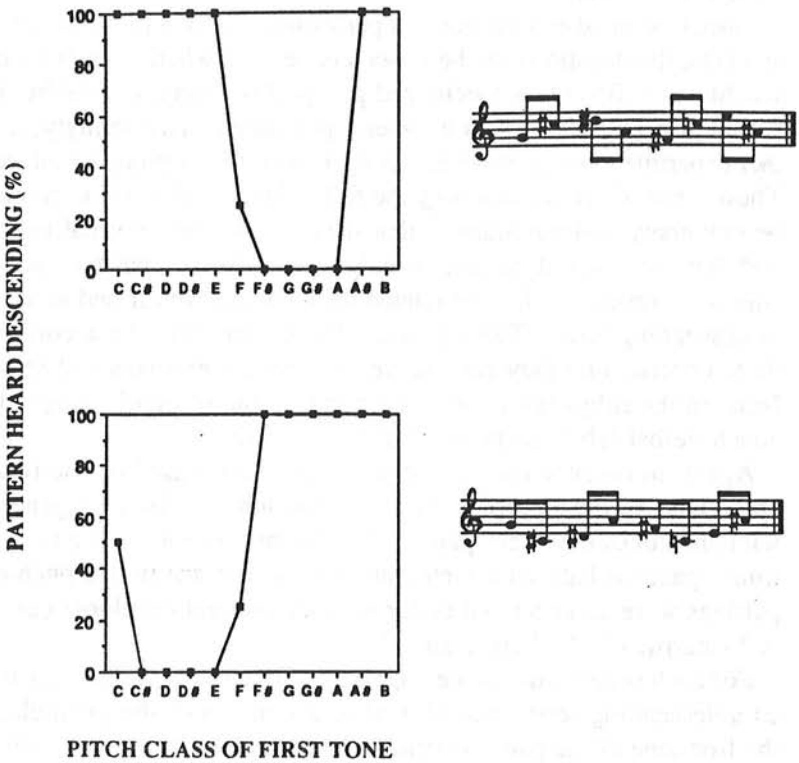

Figure 6. Graphs on left show percentages of judgments that a tone pair formed a descending series, plotted as a function of the pitch class of the first tone of the pair. Results are from the first and second subjects, for tones generated under the spectral envelope centered at F-sharp$_4$. Notations on right show how the two subjects perceived the identical series of tone pairs. (Data from Deutsch, 1986.)

the most part, when the first subject heard an ascending pattern the second subject heard a descending one, and vice versa. Thus, extended patterns composed of such tone pairs were heard by these listeners in radically different ways. An example is given on the right of the illustration, for the case of the series G– C-sharp, D-sharp – A, F-sharp – C_3 and E– A-sharp.

Since this initial experiment was performed on a few musically trained subjects, the question may be raised concerning whether such findings would be confined to a specialized group of listeners, or whether they would also be produced in a general population. Accordingly, a further experiment was performed, using a much larger group of subjects.[9] These were selected using only the following criteria: that they should be university undergraduates, that they should have normal hearing, and that they should be able to judge without error whether pairs of sine-wave tones which were related by a half-octave formed ascending or descending series. Twenty-nine subjects were selected according to these criteria, and they each served in a single experimental session. None of the subjects had absolute pitch, in the sense of being able to attach verbal labels to notes played in isolation.

Again, in order to control for possible effects based on the relative amplitudes of the sinusoidal components, tones pairs were generated with envelopes that were placed at different positions along the spectrum, spaced at half-octave intervals. In this case, the twelve pitch-class pairings were all generated under the four spectral envelopes centered at F-sharp$_5$, C_5, F-sharp$_4$, and C_4.[10]

For each subject, the percentage of judgments that a tone pair formed a descending series was plotted as a function of the pitch class of the first tone of the pair. As can be seen from the three plots in figure 7, such individual judgments were again strongly influenced by the positions of the tones along the pitch-class circle. Further, as can also be seen, the direction of this influence varied considerably across subjects.

In order to obtain an estimate of the prevalence of the effect in the subject population as a whole, the following procedure was used. First, it was determined for the scores for each subject whether the pitch-class circle could be bisected in such a way that none of the scores in the upper half of the circle was lower than any of the scores in the lower half. This criterion was fulfilled by twenty-two of the twenty-nine sub-

[9]Diana Deutsch, William L. Kuyper and Yuval Fisher, "The Tritone Paradox: Its Presence and Form of Distribution in a General Population," *Music Perception* (in press).

[10]The other sound parameters were as stated in Deutsch, "A musical paradox."

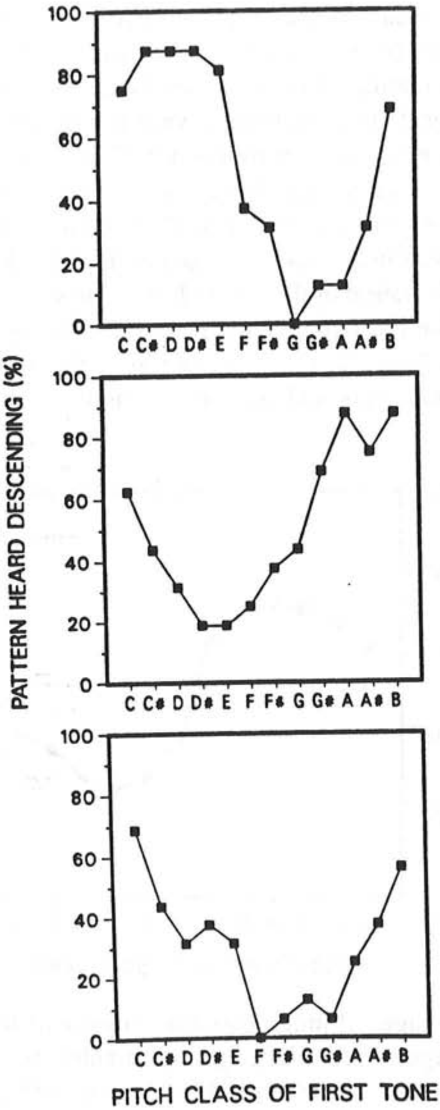

Figure 7. Percentages of judgments that a tone pair formed a descending series, plotted as a function of the pitch class of the first tone of the pair. Results from three different subjects are here displayed, averaged over four spectral envelopes. (From Deutsch, Kuyper, and Fisher, in press.)

jects. Next, in order to obtain a baseline estimate of the probability of obtaining such a result by chance, the proportion of random permutations of the scores (which could be so characterized) was determined by computer simulation. Averaged across subjects, this was found to be .027 per subject, thus yielding a vanishingly small probability of obtaining the combined result by chance. The effect was thus shown to exist to a very highly significant extent in this general population.

We may next enquire into the form of the relationship between pitch class and perceived height in this population as a whole. To examine this issue, the orientation of the pitch-class circle was normalized across subjects,[11] and the normalized data were then averaged. The resultant plot is shown in figure 8, and reveals a remarkably orderly relationship between pitch class and perceived height.

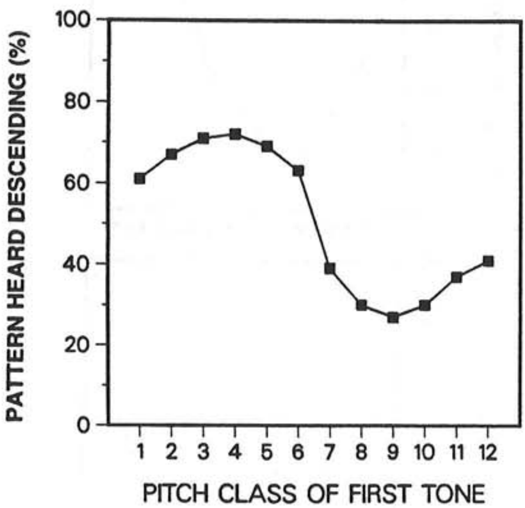

Figure 8. Percentages of judgments that a tone pair formed a descending series, averaged over a large group of subjects, with the orientation of the pitch-class circle normalized across subjects. Results were averaged over four spectral envelopes. (From Deutsch, Kuyper, and Fisher, in press.)

[11]For each subject, the pitch class circle was bisected so as to maximize the difference between the averaged scores within the two halves. The circle was then oriented so that the line of bisection was horizontal, and the data were retabulated with the leftmost pitch class of the upper half of the circle taking the first position, its clockwise neighbor taking the second position, and so on.

To examine whether the effect might be related to musical training, two analyses were performed. First, the absolute size of the effect was estimated for each subject separately by subtracting the averaged score for the lower half of the normalized circle from that for the upper half. Of the fifteen subjects who showed the larger difference on this measure, seven had had two years or less of musical training. Of the fourteen subjects who showed the smaller difference, seven also fell into this category. Indeed, the two who showed the largest difference on this measure had had no musical training whatever. For the second analysis, the proportions of "trained" and "untrained" subjects whose individual scores yielded statistically significant effects were calculated, and again no significant difference between the two groups emerged. Clearly, then, musical training is not responsible for the phenomenon.

A further study examined in detail the behavior of this phenomenon in face of variations in the position of the spectral envelope.[12] Such variations could, in principle, produce effects in two ways: first, through resultant differences in the overall heights of the patterns, and second through resultant differences in the relative amplitudes of the sinusoidal components of the tones. A further issue that was examined concerned the stability of the effect when subjects were tested over relatively long time periods.

Accordingly, twelve different spectral envelopes were employed, whose positions were spaced at intervals a quarter-octave apart, spanning altogether a three-octave range. The envelope peaks stood at A_5, F-sharp$_5$, D-sharp$_5$, C_5, A_4, F-sharp$_4$, D-sharp$_4$, C_4, A_3, F-sharp$_3$, D-sharp$_3$, and C_3. Thus, the effect of overall height could be examined by comparing the combined results from the four envelopes centered in each of the three octaves, and the effect of relative amplitude could be examined by comparing the combined results from the envelopes centered at each of the four different pitch classes.[13] Four subjects were employed in the study, and they each served in nine experimental sessions.

Figure 9 displays the results from the four subjects, in each case averaged over all twelve spectral envelopes, and over all nine experimental sessions. As can be seen, the data from each subject again revealed a highly systematic relationship between pitch class and perceived height; and again the form of this relationship varied considerably across subjects. Statistically, for some subjects the form of this relationship was

[12]Diana Deutsch, "The Tritone Paradox: Effects of Spectral Variables," *Perception and Psychophysics* 41 (1987): 563-575.

[13]The other sound parameters were as in Deutsch, "A musical paradox."

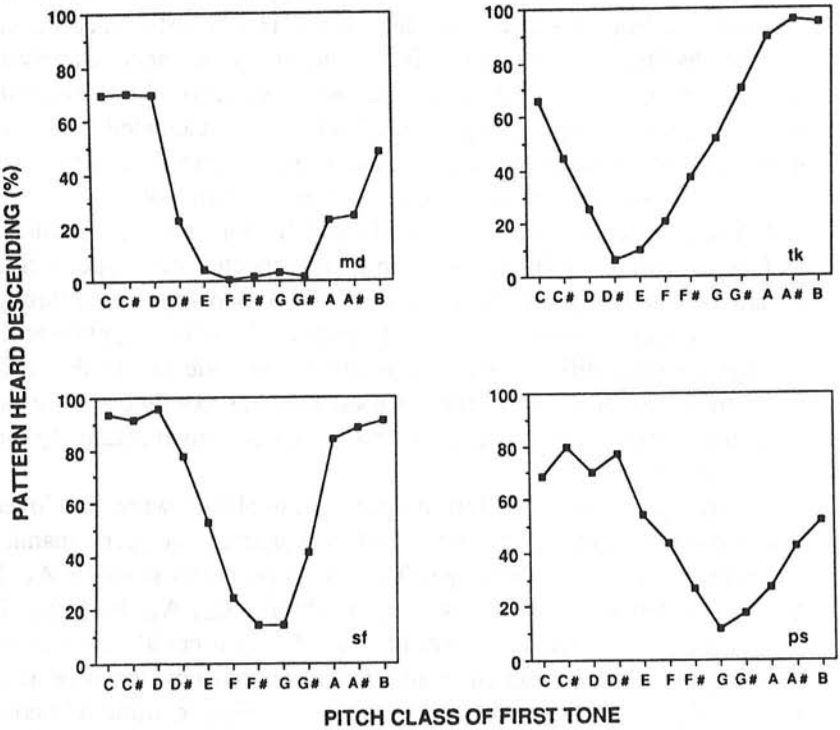

Figure 9. Percentages of judgments that a tone pair formed a descending series, plotted as a function of the pitch class of the first tone of the pair. Results from four different subjects are here displayed, averaged over twelve spectral envelopes and nine experimental sessions. (From Deutsch, 1987.)

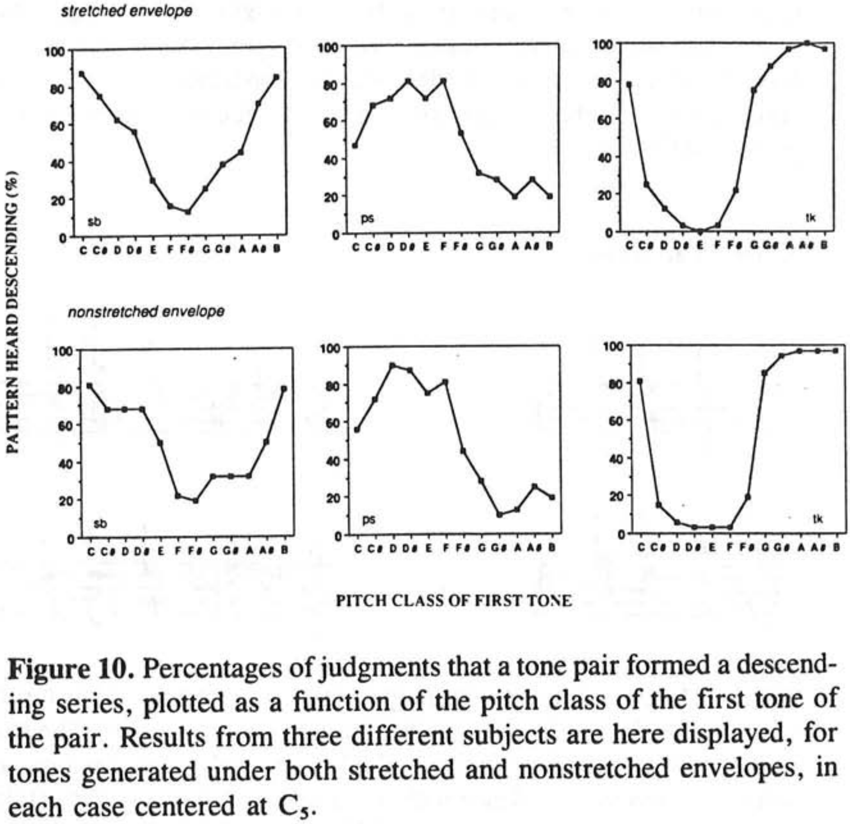

Figure 10. Percentages of judgments that a tone pair formed a descending series, plotted as a function of the pitch class of the first tone of the pair. Results from three different subjects are here displayed, for tones generated under both stretched and nonstretched envelopes, in each case centered at C_5.

influenced by the overall heights of the patterns, and for some subjects it was influenced by the relative amplitudes of the sinusoidal components of the tones; however, neither influence was necessarily present.

A further experiment considered the possible involvement of phase relationships in the effect. To examine this, tone complexes were generated whose envelopes were stretched slightly, so that the sinusoidal components stood in a ratio of 2.01:1. As a result, the phase relationships between these components were constantly varying. Figure 10 displays the results from three subjects when judging tone pairs generated under both stretched and nonstretched envelopes, both centered at C_5. It can be seen that the results under these two conditions were remarkably similar, showing that the phenomenon is not due to the processing of phase relationships.

A Two-Part Pattern

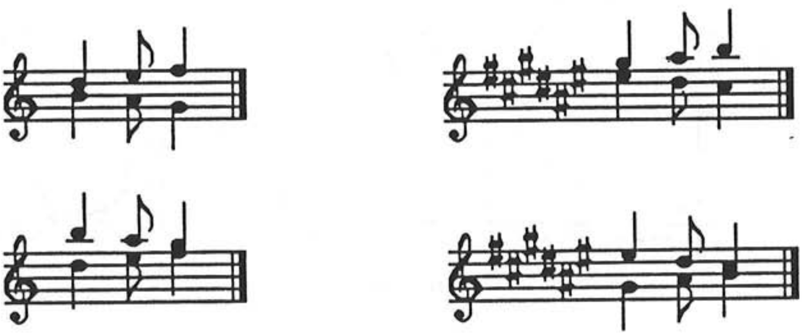

Figure 11. Configurations employed to examine the relationship between pitch class and perceived height in two-part patterns. The basic configuration was presented in both C major and F-sharp major, and is here notated in accordance with two alternative perceptual organizations. (From Deutsch, Moore, and Dolson, 1984.)

We may next ask what happens when more elaborate configurations are employed. One set of experiments[14] used the patterns shown in figure

[14]Diana Deutsch, F. Richard Moore, and Mark Dolson, "Pitch Classes Differ with Respect to Height," *Music Perception* 2 (1984): 265-271; and "The perceived height of octave-related complexes," *Journal of the Acoustical Society of Americal* 80 (1986): 1346-1353.

11. The first was in C major and consisted of the series D – E – F presented simultaneously with the series B – A – G. The second was a transposition of the first to F-sharp major, and so consisted of the series G-sharp – A-sharp – B presented simultaneously with the series E-sharp – D-sharp – C-sharp. The patterns were perceptually organized as two melodic lines in accordance with pitch proximity, so that the listener heard one line that ascended by a minor third, together with another line that descended by a major third. However, some listeners heard the ascending line as higher and the descending line as lower whereas other listeners heard the descending line as higher and the ascending line as lower. These two different perceptual organizations are shown in figure 11.

In one experiment,[15] the two patterns were each generated under envelopes that were placed at six different positions along the spectrum. These were spaced at half-octave intervals, so that their peaks spanned a two and one-half octave range. The envelopes were centered at C_6, F-sharp$_5$, C_5, F-sharp$_4$, C_4, and F-sharp$_3$. Figure 12 shows, as examples, the spectral compositions of the tones comprising the chord D/B, generated under each of these envelopes. On each trial, one of the patterns was presented, and subjects judged whether the higher of the two lines formed an ascending or a descending series; from these judgments it was inferred which tones were heard as higher and which as lower.[16]

Analogous effects were also found to occur with this type of pattern. When played in one key it was heard with the higher line ascending, yet when played in a different key it was heard with the higher line descending. When the pattern was transposed, the relative heights of the different pitch classes were preserved, so that a perceived interchange of voices resulted. Further, when the pattern was played in any one key, listeners differed radically in terms of which line they heard as higher and which as lower.

Table 1 shows, for both groups of subjects, the percentages of judgments where the higher line formed an ascending series, as a function both of key and also of position of the spectral envelope. It can be seen that for both groups of subjects, judgments depended almost entirely on the key in which the pattern was presented. Thus Type-A

[15]Deutsch, Moore, and Dolson, "The perceived height..."

[16]In all patterns, the first and third tones within a series were 500 msec in duration, and the second tone was 250 msec in duration. There were no pauses between tones within a pattern. On each trial, one of the patterns was presented three times in succession, the presentations being separated by 750 msec pauses. Trials were in blocks of twenty-four, with ten- second pauses between trials within a block, and five-minute pauses between blocks. The patterns within a block were presented in random order.

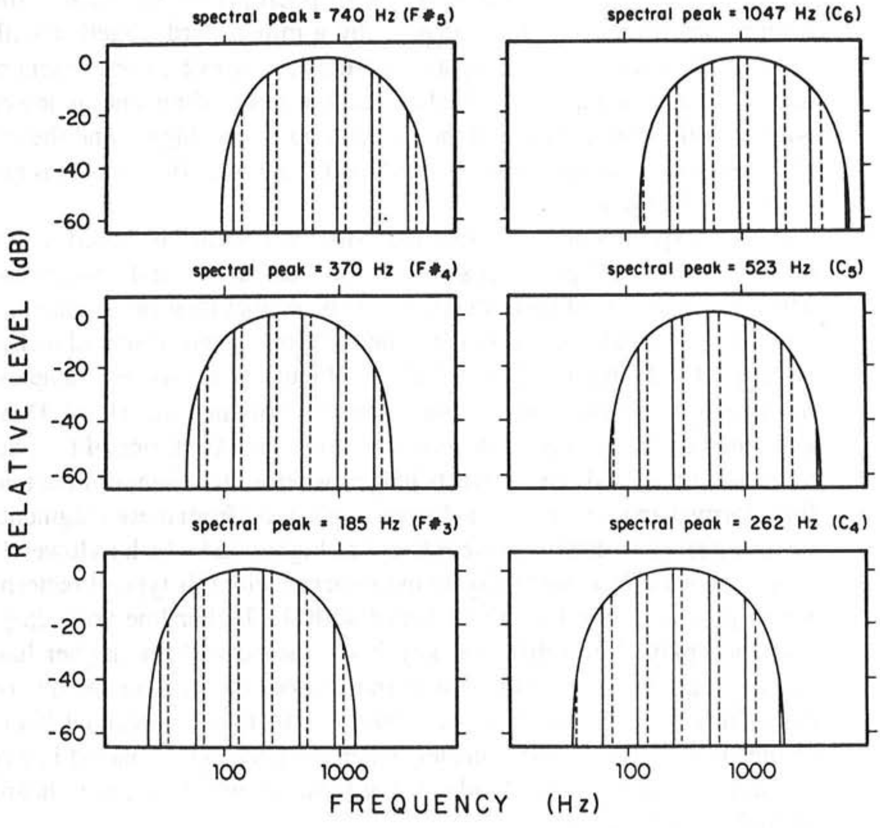

Figure 12. Representation of the spectral compositions of tones employed to examine the influence of pitch class on perceived height in two-part patterns. The chord B/D is here displayed, generated under six spectral envelopes. Solid lines indicate tones of pitch-class B, and dashed lines tones of pitch-class D. (From Deutsch, Moore, and Dolson, 1986.)

Table 1. Percentages of judgments that the higher line formed an ascending series. Tones were generated under nonstretched envelopes. (From Deutsch, Moore, and Dolson, 1986.)

	Spectral peak						
	F-sharp$_3$	C_4	F-sharp$_4$	C_5	F-sharp$_5$	C_6	
C major	100	100	100	97	94	100	Type-A subjects
F-sharp major	0	0	0	0	0	0	

	Spectral peak						
	F-sharp$_3$	C_4	F-sharp$_4$	C_5	F-sharp$_5$	C_6	
C major	3	9	0	3	6	0	Type-B subjects
F-sharp major	97	97	97	97	97	100	

C-major pattern: ascending line composed of pitch classes D, E, and F; descending line composed of pitch classes B, A, and G.
F-sharp major pattern: ascending line composed of pitch classes G-sharp, A-sharp, and B; descending line composed of pitch classes E-sharp, D-sharp, and C-sharp.

subjects heard the C major pattern with notes D, E, and F as higher and B, A, and G as lower, for all positions of the spectral envelope. They also they heard the F-sharp major pattern with notes E-sharp, D-sharp, and C-sharp as higher and notes G-sharp, A-sharp, and B as lower, for all positions of the spectral envelope. Figure 13 shows, on the left, the pitch-class circle oriented with respect to height so as to reflect this pattern of results. Type-B subjects, on the other hand, produced results which were the converse of those of Type-A. They consistently heard the C-major pattern with notes B, A, and G as higher and notes D, E, and F as lower, for all positions of the spectral envelope. They also heard the F-sharp-major pattern with notes G-sharp, A-sharp, and B as higher and notes E-sharp, D-sharp, and C-sharp as lower, for all positions of the spectral envelope. Figure 13 shows, on the right, the pitch-class circle oriented with respect to height so as to reflect these results.

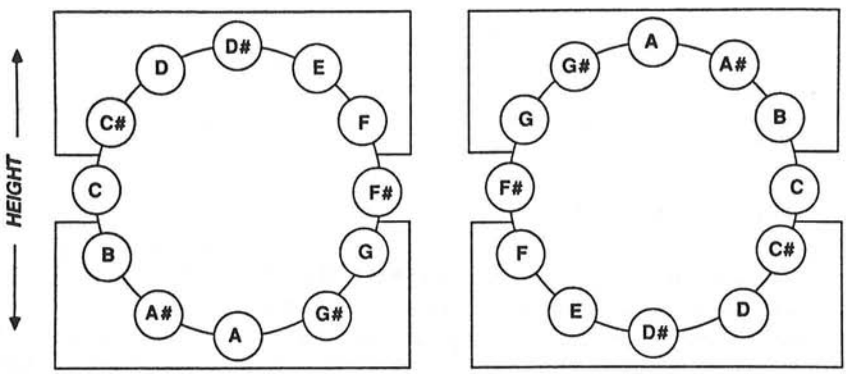

Figure 13. Two orientations of the pitch class circle with respect to height, as reflected in the judgments of Type-A subjects (shown on left), and Type-B subjects (shown on right). (From Deutsch, Moore, and Dolson, 1986.)

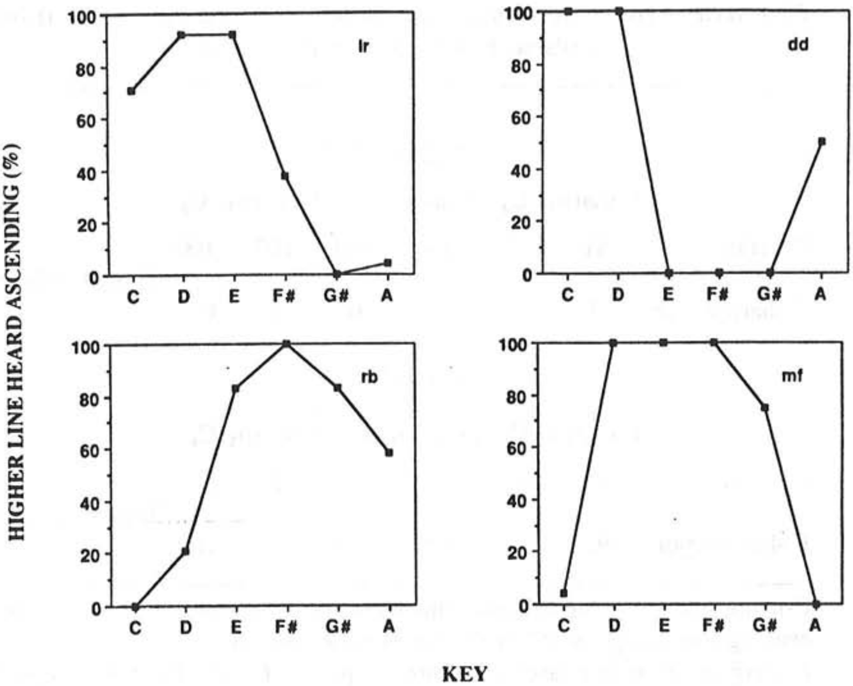

Figure 14. Percentages of judgments that the higher line of the two-part pattern formed an ascending series, as a function of the key in which the pattern was presented. The results from four different subjects are here displayed, averaged over three spectral envelopes.

In a further experiment,[17] tone complexes were generated under envelopes that were stretched slightly, so that the sinusoidal components stood in a ratio of 2.01:1. As a result, the phase relationships between these components were constantly varying. As shown in table 2, the influence of pitch class on perceived height was found to be unaffected by this manipulation.

Table 2. Percentages of judgments that the higher line formed an ascending series. Tones were generated under stretched envelopes. (From Deutsch, Moore, and Dolson, 1986.)

	Spectral peak						
	F-sharp$_3$	C$_4$	F-sharp$_4$	C$_5$	F-sharp$_5$	C$_6$	
C major	91	97	100	100	100	100	Type-A subjects
F-sharp major	3	3	3	0	0	0	
	Spectral peak						
	F-sharp$_3$	C$_4$	F-sharp$_4$	C$_5$	F-sharp$_5$	C$_6$	
C major	13	6	0	3	3	3	Type-B subjects
F-sharp major	97	97	100	94	97	100	

C-major pattern: ascending line composed of pitch classes D, E, and F; descending line composed of pitch classes B, A, and G.
F-sharp major pattern: ascending line composed of pitch classes G-sharp, A-sharp, and B; descending line composed of pitch classes E-sharp, D-sharp, and C-sharp.

In order to examine this phenomenon in greater detail, I generated the pattern in six different keys; namely C, D, E, F-sharp, G-sharp, and A-sharp. The results from four subjects are shown in figure 14, averaged over the spectral envelopes centered at F-sharp$_5$, C$_5$, and F-sharp$_4$. It can be seen that all subjects showed highly systematic effects of key, and they also differed considerably in terms of the direction in which key influenced their judgments.

Figure 15 illustrates these points more closely, by taking as an ex-

[17] Deutsch, Moore, and Dolson, "The perceived height...."

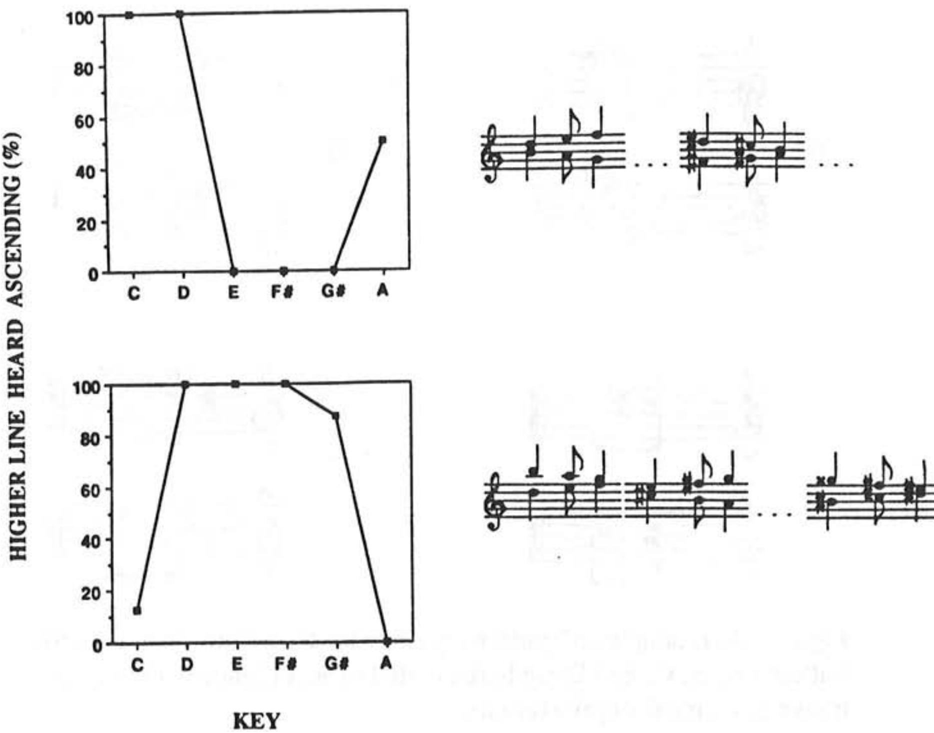

Figure 15. Graphs on left display the percentages of judgments that the higher line of the two-part pattern formed an ascending series, as a function of the key in which the pattern was presented. The results from two subjects are displayed, for patterns generated under the spectral envelope centered at C_5. Notations on right show how the identical patterns were perceived by the two subjects.

ample the judgments of two of the subjects for tones generated under the envelope centered at C_5. So taking the first subject and moving from left to right, the pattern in the key of C was heard with the higher line ascending, as it was in the key of D. But in the key of E it was heard with the higher line descending; and so on. It can be seen that the second subject produced judgments which were virtually the converse of the first. The perceived heights of tones in this pattern were therefore found to vary systematically with their positions along the pitch-class circle, in a fashion analogous to the tritone paradox.

The Semitone Paradox

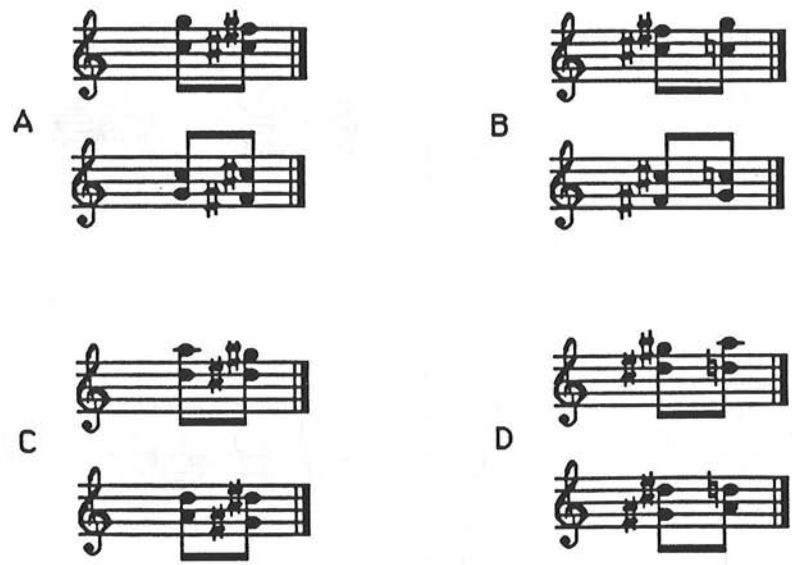

Figure 16. Examples of patterns giving rise to the semitone paradox. Patterns A, B, C, and D are here notated in accordance with two alternative perceptual orgainizations.

The two-part pattern just described consisted of melodic lines which encompassed a relatively large region of the pitch-class circle. This could, in principle, give the subject freedom to focus on different parts of the pitch-class circle in making his judgments, which creates an ambiguity of interpretation. So in order to examine the phenomenon in a more fine-grained fashion, a different two-part pattern was devised.[18] As shown by the examples in figure 16, this pattern comprised two simultaneous tone pairs, one of which ascended by a semitone and the other of which descended. These tone pairs were diametrically opposed along the pitch class circle. The listener, following the principle of proximity, perceived these patterns as two stepwise lines which moved in contrary motion. However, as shown in figure 17, proximity could not here be used as a cue to determine which line would be perceived as higher and which as lower.

Subjects were presented with such patterns in random order, so that

[18]Diana Deutsch, "The semitone paradox," (in preparation).

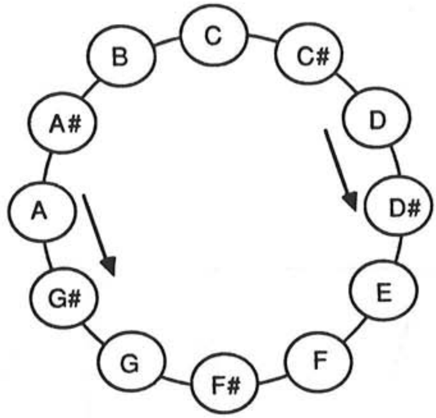

Figure 17. Representation of a pattern giving rise to the semitone paradox, in terms of relationships within the pitch-class circle. The pattern notated in figure 16C is here depicted.

each pitch class served equally often as the first tone of an ascending pair, and also as the first tone of a descending pair. The following pitch class combinations were therefore employed: C – C-sharp/G – F-sharp; C-sharp – D/G-sharp – G; D – D-sharp/A – G-sharp; D-sharp – E/A-sharp – A; E – F/ B – A-sharp; F – F-sharp/C – B; F-sharp – G/C-sharp – C; G – G-sharp/D – C-sharp; G-sharp – A/D-sharp – D; A – A-sharp/ E – D-sharp; A-sharp – B/F – E; and B – C/F-sharp – F. Subjects judged on each trial whether the higher line formed an ascending or a descending series; from these judgments it was inferred which tones were heard as higher and which as lower.[19]

The patterns were generated under envelopes which were placed at twelve different positions along the spectrum, which were spaced at quarter-octave intervals, spanning a three-octave range. The envelope peaks stood at A_5, F-sharp$_5$, D-sharp$_5$, C_5, A_4, F-sharp$_4$, D-sharp$_4$, C_4, A_3, F-sharp$_3$, D-sharp$_3$, and C_3. Four subjects were employed in the study, and each served in nine experimental sessions.

Figure 18 shows the results from the four subjects, in each case averaged over all twelve spectral envelopes and all nine experimental sessions. A strong influence of pitch class on perceived height was again

[19]All tones were 500 msec in duration, and there were no pauses between the tones within a pattern. Trials were presented in blocks of twelve, with five-second pauses between trials within a block, and one-minute pauses between blocks. Each block consisted of tones generated under one of the twelve envelopes and contained one example of each of the twelve pitch-class combinations.

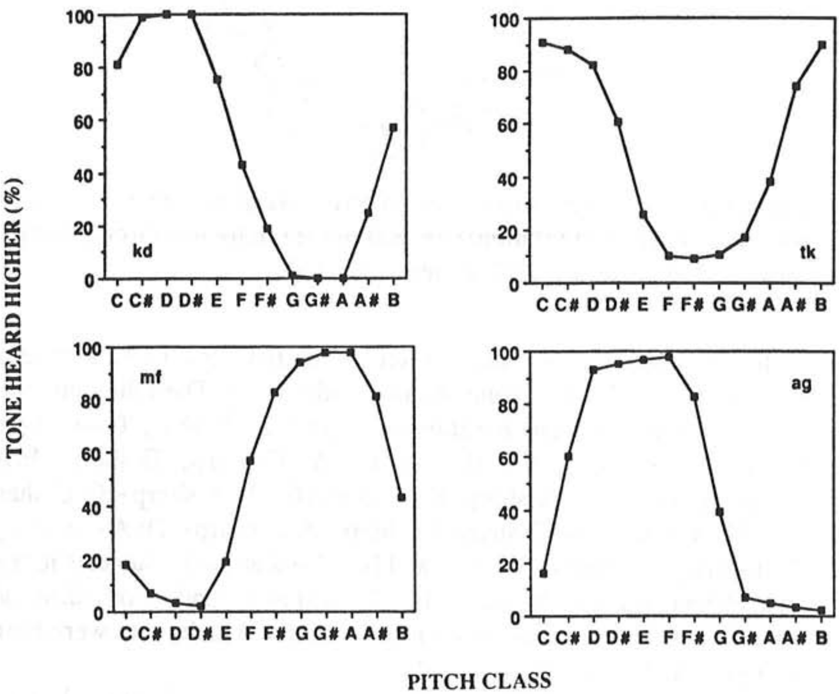

Figure 18. The semitone paradox. Percentages of trials in which a tone was heard as part of the higher line, plotted as a function of the pitch class of the tone. Results from four subjects are here displayed, averaged over twelve spectral envelopes and nine experimental sessions.

apparent, and this was also found to be robust in face of variations in both the overall heights of the patterns and also with reference to the relative amplitudes of the sinusoidal components of the tones. Further, there were again striking differences between the subjects in the direction of the relationship between pitch class and perceived height.

Figure 19 illustrates these points more closely by taking as an example the judgments of two of the subjects, in this case with the spectral envelope centered on C_5. Here the series of pitch class combinations D – D-sharp/A – G-sharp, C-sharp – C/F-sharp – G and D-sharp – E/A-sharp – A were presented, and it can be seen that the subjects heard this extended pattern in ways that were radically different from each other. So the same perceptual principles were here manifest as for the other types of pattern.

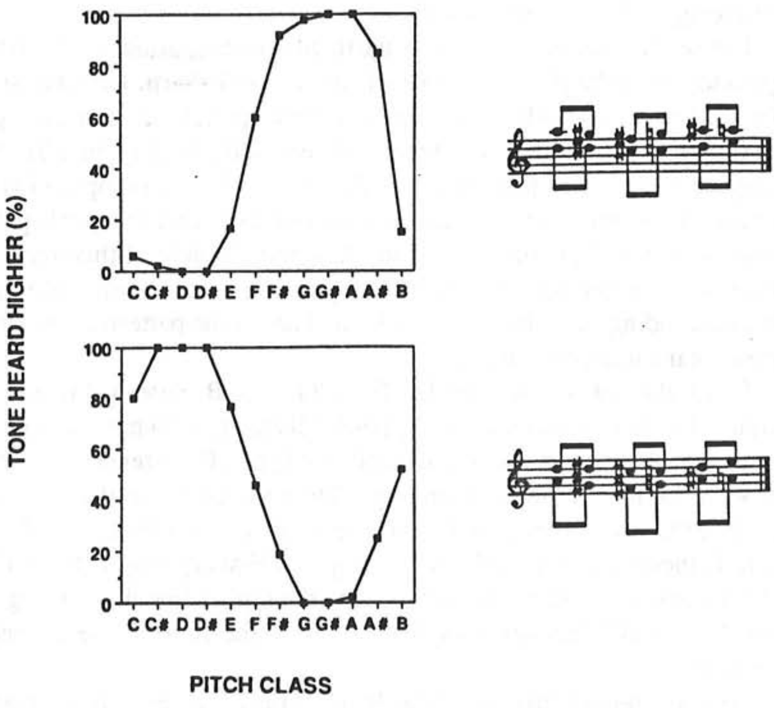

Figure 19. The semitone paradox. Graphs on left display the percentages of trials in which a tone was heard as part of the higher line, plotted as a function of the pitch class of the tone. Results from two subjects are here displayed, for patterns generated under the spectral envelope centered at C_5. Notations on right show how the identical series of patterns was perceived by the two subjects.

DISCUSSION

The phenomena reported here are highly unexpected on a number of grounds. First, they provide clear counter-examples to the principle of invariance under transposition—a principle which has generally been regarded as self-evident. In the case of the tritone paradox, transposing the pattern resulted in a perceived inversion; in the case of the two-part patterns, transposition resulted in a perceived interchange of voices.

Particularly convincing demonstrations of these phenomena may be produced by tape-recording examples of such patterns and then playing the tape back at different speeds. This manipulation shifts the spectra of the patterns up or down in log frequency, so that different pitches are produced. One would assume that the patterns would, as a result, simply be heard as transposed; however, they are heard as having radically changed their shape as well.

Let us take as an example a particular instantiation of the tritone paradox—say the pitch-class combination D – G-sharp. Let us also take two listeners, one who hears this particular pattern as descending (as in figure 4) and another who hears it as ascending (as in figure 5). After playing the pattern at normal speed, we speed the tape up so that the entire spectrum is transposed up a half-octave, and the pitch pattern now becomes G-sharp – D instead. As a result solely of this manipulation, the listener who had heard the pattern as descending now hears it as ascending, and the listener who had heard the pattern as ascending now hears it as descending!

Similarly, let us take the D – E – F/G – A – B pattern illustrated in figure 11. Let us also take a "Type-A" listener, who hears the pattern with the higher line ascending, and a "Type-B" listener, who hears it with the higher line descending. When the tape speed is increased so that the entire spectrum is transposed up a half-octave, and the pattern is thereby transposed from C major to F-sharp major, the "Type-A" listener now hears the pattern with the higher line descending, and the "Type-B" listener now hears it with the higher line ascending instead!

Since the relative heights of the different pitch classes remain invariant in face of overall shifts in the frequency spectrum, the findings demonstrate that pitch class and perceived height are not orthogonal dimensions, as had been supposed. The findings are therefore at variance with the helical model of pitch described at the beginning of the chapter. This point is illustrated with reference to figure 20. Take, for example, the listener who hears note A as higher and note D-sharp as lower,

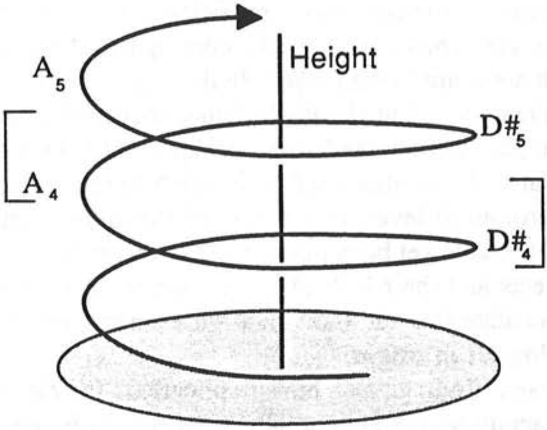

Figure 20. The relative heights of tones in different spectral regions, according to the model of pitch as a geometrically regular helix.

regardless of the position of the spectral envelope. The bracket on the right indicates a region of height which corresponds to a spectral envelope centered on F-sharp$_4$. Within this region, as shown in the diagram, the note A should indeed be heard as higher and the note D-sharp as lower. However, consider now the region of height indicated by the bracket on the left, and which corresponds to a spectral envelope centered on C$_5$. Here, according to the model, D-sharp should be heard as higher and A as lower instead. *However, the listener continues to hear A as higher and D-sharp as lower, regardless of the spectral region the tones are in*. We can see that the model of pitch as a geometrically regular helix cannot accommodate this phenomenon.

Another highly unexpected aspect of these findings concerns the pronounced differences between listeners in how such patterns are perceived. It is particularly striking that these differences occur as strongly among those who are musically proficient as among those who are musically naive. In general, it is assumed that while people differ in terms of how they might group or classify musical materials, or how accurately they can describe them, the issue of whether a simple two-note pattern forms an ascending or a descending series should not be a matter of dispute. These findings therefore lead us to wonder what other differences between listeners might exist that have not yet been uncovered. Such differences might give rise to disagreements about music which are at present considered to be aesthetic or otherwise evaluative in nature.

Variations between listeners have also been found to occur in the octave and scale illusions.[20] Here, proficient musicians may differ strikingly as to where notes appear to be coming from, and even how many and which notes are being played. In the case of these illusions, strong correlates have been found with the handedness of the listener, and even with his or her familial handedness background.[21] One might therefore assume that such variations reflect innate differences between listeners at the neurological level. In the case of the present set of paradoxes, no correlates have yet been obtained, though the highly orderly nature of the effects and their lack of association with musical training lead one to speculate that the individual differences found here are likely to be biological in origin.

The present findings also have implications for theories of absolute pitch, a faculty which is generally supposed to be confined to a few rare individuals. In the present experiments, where judgments of height were influenced by the positions of the tones along the pitch-class circle, listeners were, in effect, using absolute pitch in making these judgments. Yet, as described, this influence of pitch class was shown to exist to a highly significant extent in a general population.[22] Given this evident ability to employ absolute pitch indirectly,[23] it is puzzling that people do not in general possess this faculty, in the sense of being able to attach verbal labels to notes presented out of context. One is reminded here of the syndrome of color anomia,[24] in which people may perform normally on nonverbal color tasks but are unable to name colors when these are presented in isolation. Although color anomia is very rare, perhaps the majority of us (i.e., those who do not possess absolute pitch as conventionally defined) have an analogous syndrome with regard to pitch class.

The extent to which effects such as described here occur in other contexts remains to be investigated. However, recent work by the author

[20]Diana Deutsch, "An auditory illusion," *Nature* 251 (1974): 307-309; and also "Two-channel listening..." and "Musical illusions."

[21]Diana Deutsch, "The Octave Illusion in Relation to Handedness and Familial Handedness Background," *Neuropsychologia* 21 (1983): 289-293.

[22]Deutsch, Kuyper and Fisher, "The Tritone Paradox: Its Presence...."

[23]A related point has been made concerning key identification; see Ernst Terhardt and W. Dixon Ward, "Recognition of Musical Key: Exploratory Study," *Journal of the Acoustical Society of Americal* 72 (1982): 26-33; and Ernst Terhardt and Manfred Seewann, "Aural Key Identification and Its Relationship to Absolute Pitch," *Music Perception* 1 (1983): 63-83. These authors found that musicians could judge whether or not a passage was played in the correct key, even though most of the subjects claimed not to have absolute pitch. Such judgments were even made when differences as small as a semitone were at issue.

[24]Norman Geschwind and M. Fusillo, "Color-Naming Defects in Association with Alexia," *Archives of Neurology* 15 (1966): 137-146.

has shown that the influence of pitch class on perceived height persists when the tones are subjected to a number of time-varying manipulations, such as superimposing a vibrato, a tremolo, or a fast decay such as to produce the impression of a plucked string. The effect has also been found to persist when the sinusoidal components within each tone complex are all replaced with sawtooth waves, so that each component is replaced by a set of tones comprising a harmonic series. Similarly, the effect persists when the sinusoids are replaced with square waves, so that each component is replaced by the odd-numbered partials of a harmonic series. The spectra of these more elaborate tone complexes are quite similar to those produced by a group of natural instruments playing simultaneously, with their fundamental frequencies standing in octave relation. Therefore, good reason exists to expect that analogous effects should be obtainable from an ensemble of natural instruments playing at appropriate loudness levels relative to each other.

More generally, we may speculate that effects such as shown here might occur, at least to some extent, in listening to music played by natural instruments under conditions giving rise to registral ambiguities. Consider, for example, orchestral contexts in which instruments of different types play with multiple octave doublings. Under such conditions the perceived heights of tones are often not clearly defined. The question of how musical patterns are really heard under such circumstances has not yet been the subject of formal investigation.

At a strictly theoretical level, the phenomena described in this essay lead us to question the relationship between a musical note as it is written in a score, and as depicted in alphanumeric form, or as it is perceived. It is generally taken for granted that a clear correspondence between the two must surely exist.[25] However, the present work shows that under certain conditions at least the assumption is incorrect—if only because listeners disagree among themselves as to which notes they hear. It would appear that a symbolic description, which was originally developed as a practical convenience (for example, as instructions to the performer),

[25]It is often assumed that for single notes composed of harmonically related partials, the perceived pitch corresponds to that of the fundamental; see, for example, J. F. Schouten, "The Residue and the Mechanism of Hearing," *Proceedings of the Koninklgke Nederlandse Akademie van Wetenshappen* 43 (1940): 991-991. However, listeners may produce pitch matches to frequencies other than the fundamental even with strictly harmonic tone complexes (see, for example, Ernst Terhardt, "Gestalt principles and music perception," in *Auditory Processing of Complex Sounds*, ed. William A. Yost and Charles S. Watson (Hillsdale, 1987): 157-167. In addition, compelling demonstrations of the failure of residue pitch have been presented by John R. Pierce, "What Do *We* Hear?" in a paper presented at the *81st Audio Engineering Society Convention*, Los Angeles, 1986.

has been elevated through common usage, and without logical or empirical support, to the status of a fundamental theoretical construct.[26]

In conclusion, the findings described in this essay cannot be readily accommodated within present theory, and raise considerably more questions than they answer. However, as Leonard B. Meyer, who has always argued so eloquently for the usefulness of the empirical approach to music, wrote: "Were complete information and incontrovertible theory a prerequisite for understanding, science, for example, would never have even begun."[27]

[26]Eugene Narmour, in *Beyond Schenkerism: The Need for Alternatives in Music Analysis* (Chicago, 1977) has advanced analogous arguments concerning the logical status of the *Ursatz*.

[27]Leonard B. Meyer, *Explaining Music*, p.14.

[28]I am indebted to a number of people for their collaboration and assistance in this research, specifically to F. Richard Moore and Mark Dolson for software used in synthesizing the tones, and to Lee Ray, Yuval Fisher, and William L. Kuyper for assistance and collaboration in various phases of the study. The work was supported by grants from the System Development Foundation, and from the UCSD Biomedical Research Support Program.