Paradoxes of Musical Pitch

Certain series of tones appear to ascend or descend infinitely in pitch. Other patterns change when shifted in key and indicate an influence of speech on the perception of music

by Diana Deutsch

The endless staircase stands as a classic visual paradox, continuously tricking the eye into a geometrically impossible journey. First devised in the 1950s by L. S. Penrose and Roger Penrose of the University of London and later made famous by the Dutch artist M. C. Escher, this paradox has a rich set of acoustic counterparts. In the early 1960s Roger N. Shepard of Bell Telephone Laboratories produced a rather remarkable example. He repeatedly played the same sequence of computer-generated tones that moved up in an octave. Instead of hearing the pattern stop and then start again, listeners heard the pattern ascend endlessly in pitch. When Shepard reversed the direction, the subjects heard the pattern descend endlessly.

Such inquiries amount to far more than an acoustic form of aesthetic diversion. Research into the way individuals hear particular sequences of tones reveals how the brain uses different cues to make sense of ambiguous sounds. Indeed, the latest studies suggest that perception of certain musical paradoxes is related to the processing of speech. It appears that during childhood individuals gradually acquire a representation of pitch that is peculiar to a particular language or dialect. Hence, a native of California will perceive a certain pattern of tones differently from a native of the south of England. Such studies have revealed that a common influence on the perception of both speech and musical pitch exists in individuals.

DIANA DEUTSCH is professor of psychology at the University of California, San Diego, where she received her Ph.D. She obtained a first-class honors B.A. degree from the University of Oxford. Her work has centered on the perception of sound and music, including ways in which sound patterns are analyzed and represented in memory, as well as the neurological underpinnings of these processes. This is her second article for *Scientific American.*

Furthermore, the research has overturned some long-standing assumptions, particularly one concerning perceptual equivalence for musical patterns. This assumption states that a musical passage remains identifiable even if it is presented in a key different from that in which it was originally heard. But on the contrary, certain pitch paradoxes show that this principle is not universal. Rather the brain may completely reinterpret the relations between tones transposed to another key. This notion is as paradoxical as the idea that a visual shape might undergo a metamorphosis if shifted to a different location in space.

Research that supports these and other conclusions has deep historical roots in studies of the physics of music and sound. Indeed, the physical basis of musical pitch has fascinated scientists since antiquity. Pythagoras established that the pitch of a vibrating string varies inversely with its length: the shorter the string, the higher the pitch. He also demonstrated that certain musical intervals—the pitch relation between two tones—correspond to ratios formed by different lengths of string. In the 17th century Galileo and the French mathematician and theologian Marin Mersenne showed that the basis of these associations lay in the relation between string length and frequency of vibration.

Mersenne also demonstrated the existence of overtones, or harmonics, in vibrating bodies. That is, a vibration occurs both at the frequency corresponding to the perceived pitch (the fundamental frequency) and at frequencies that are whole number multiples of the fundamental (harmonics). In other words, a tone whose fundamental frequency is 100 hertz contains components at 200 hertz, 300 hertz, 400 hertz and so on. In the 1930s Jan Schouten of Philips Laboratory in Eindhoven showed that the auditory system exploits this phenomenon. When presented with a harmonic series, we can perceive a pitch that corresponds to the fundamental frequency, even if the fundamental itself is missing.

The relations between pitches enable us to hear musical patterns. When two tones are presented simultaneously or in succession, we perceive a musical interval. Intervals are heard to be the same in size when the fundamental frequencies of their component tones stand in the same ratio. (Technically, the tones within each pair are separated by the same distance in log frequency.)

This principle forms one of the bases of the traditional musical scale. The smallest unit of this scale is the semitone, which is the pitch relation formed by two adjacent notes on a keyboard. The semitone corresponds to a frequency ratio of approximately 18:17. Intervals composed of the same number of semitones are given the same name. For example, the interval corresponding to a ratio of 2:1 (12 semitones) is termed an octave, the interval corresponding to a ratio of 3:2 (seven semitones) is termed a fifth and the ratio 4:3 (five semitones) is called a fourth.

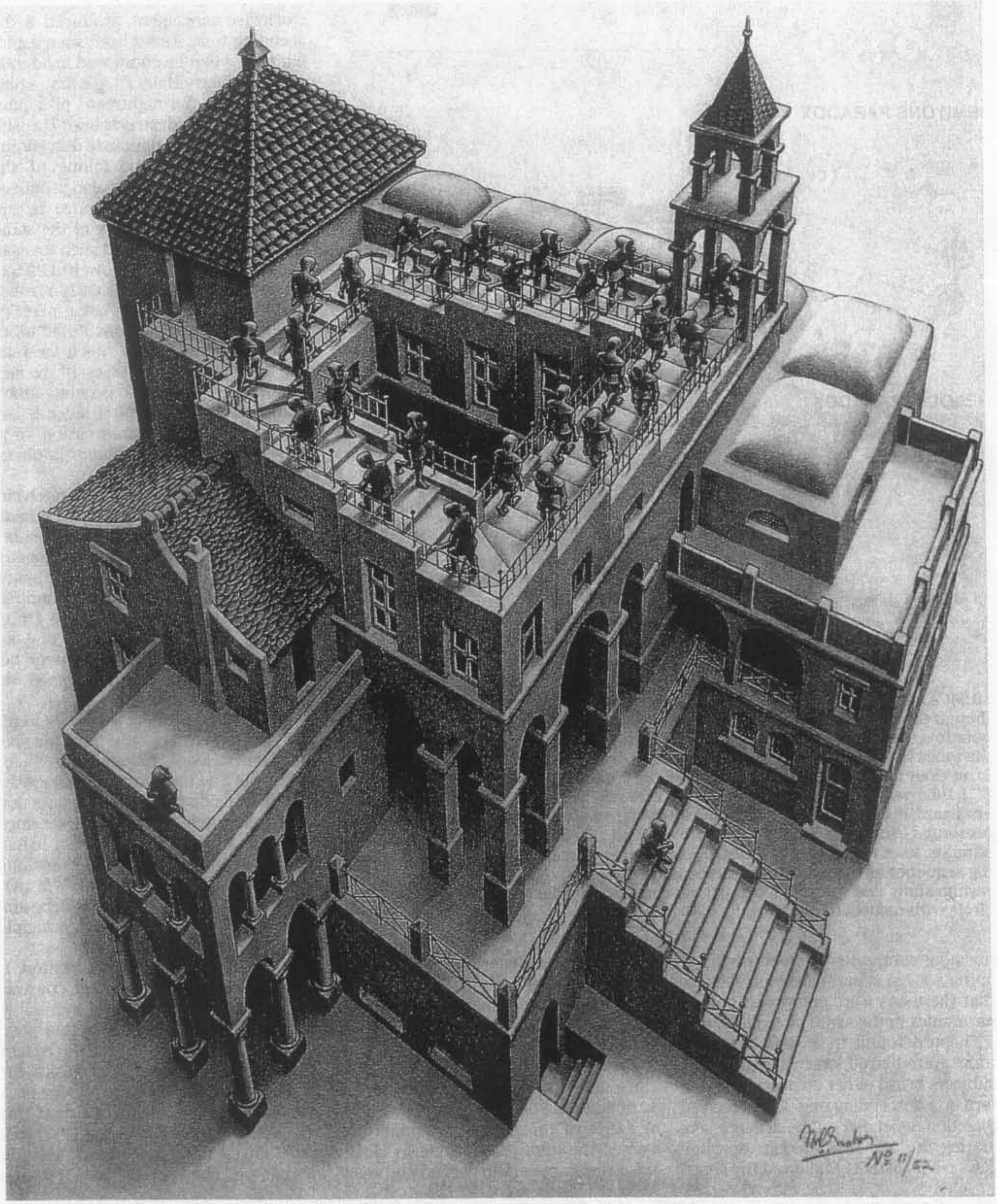
Tones related by octaves are in a sense perceptually equivalent. Each of the 12 semitones in an octave is assigned a name (C, C#, D and so on). The entire scale (called the chromatic scale) consists of the repetitive occurrence of this series of note names across octaves. The note names are identified by subscripts. For example, middle C can be written as C_4. The C one octave lower is C_3, and the one above is C_5.

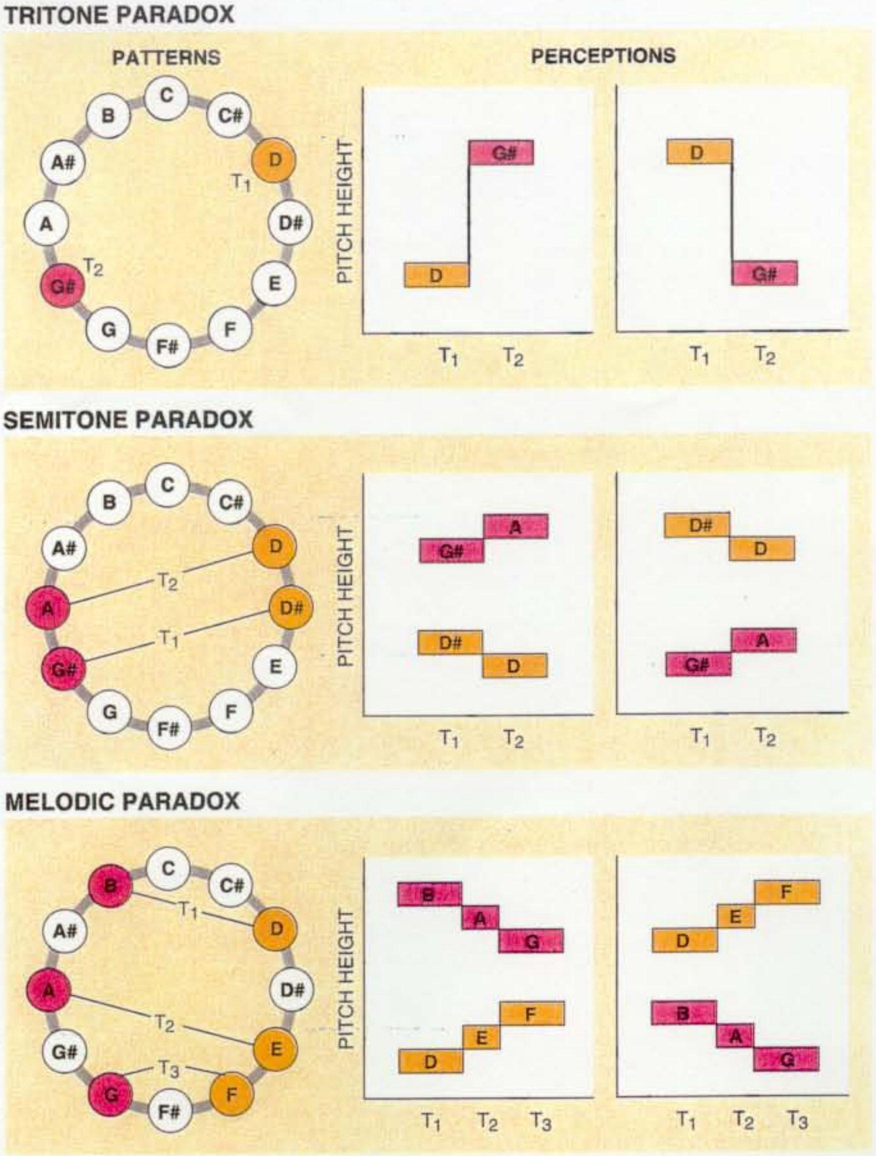
The pitch of a tone can thus be regarded as varying along two dimensions. The first, known as pitch height,

***ASCENDING AND DESCENDING,* a lithograph by M. C. Escher, visually parallels the musical illusion of tones that appear to rise or fall endlessly in pitch.**

extends from low to high, which we can experience by sweeping a hand all the way up a keyboard. The second is the circular dimension of pitch class, which defines a tone's position within the octave. Researchers refer to this dimension as the pitch class circle. The circle leads to an immediate assumption: it is nonsensical to ask whether one tone, say, C, is higher than another, such as F#. To clarify the question, one would need to give the octaves in which the two tones occur.

In the absence of such information the human brain still tries to organize tones so that it can judge relative pitch. Shepard demonstrated this phenomenon in 1964. Using a music synthesis program developed by his colleague Max Mathews, he generated a series of tones that were clearly defined in terms of pitch class but in which the octave containing the tones was unclear. Each tone consisted of a set of si-





MUSICAL PARADOXES occur when sequences of tones appear to rise or fall even though the tones lack the physical cues normally used to judge pitch height. Such paradoxes can be understood in terms of the pitch class circle, which represents the tones in an octave. The tones played are opposite one another along the circle. In an example of the phenomenon, called the tritone paradox, D is played at time 1 (T_1), followed by G# at time 2 (T_2). Some listeners heard the sequence ascend; others heard it descend. In a variation called the semitone paradox, D# and G# are presented simultaneously, followed by D and A. Another version, the melodic paradox, uses three pairs of tones. In these cases, some subjects heard the ascending sequence as higher than the descending one, and others heard it as lower. The results show that the subjects must have preferred orientations of the pitch class circle with respect to pitch height.

nusoidal components (smoothly oscillating waves) separated by octaves, so that the tones were composed only of harmonics in the same pitch class.

Shepard found that when two such tones were played, one after the other, subjects heard either an ascending pattern or a descending one. The perceived direction depended on the distance separating the two tones along the pitch class circle: listeners followed the shorter distance between the tones. For example, subjects heard the pair C#-D as ascending, because the shorter distance here is clockwise. Analogously, the pair A-G# was always heard as descending.

This finding enabled Shepard to produce the striking demonstration described at the beginning of this article. A series of tones that repeatedly traverses the pitch class circle in clockwise steps appears to ascend endlessly in pitch. If the series of tones traverses the circle in counterclockwise steps, it appears to descend infinitely.

Jean-Claude Risset, now at the Laboratory for Mechanics and Acoustics at the CNRS in Marseilles, produced an intriguing variant. He created a single tone that glided around the pitch class circle in a clockwise direction. The tone appeared to ascend endlessly. Counterclockwise movement produced a descending tone. Risset used such a gliding tone when he composed incidental music to Pierre Halet's *Little Boy*, a play that depicts the nightmare of a pilot involved in the destruction of Hiroshima. A tone of continuously descending pitch symbolized the falling of the atomic bomb. Risset has also produced a gliding tone that appears to ascend and to descend endlessly at the same time. He has incorporated such tones in orchestral works, with powerful effect.

In my laboratory I recently created pitch circularity effects using a set of tones, each of which constituted a full harmonic series but in which the relative amplitudes (loudnesses) of the harmonics were such as to generate ambiguities of perceived pitch height. Listeners obtained an impression of a series that ascended infinitely in pitch.

These demonstrations of pitch circularity illustrate that the human mind tends to form linkages between elements that are close together rather than those that are far apart. Analogous phenomena can be found in vision. For example, we tend to group together dots that are next to one another and to perceive movement between neighboring lights turned on and off in succession.

What happens, then, when two tones are related by exactly half an octave, such as C followed by F# or G# by D? The tone pairs are separated by the same distance in either direction along the pitch class circle. Over such an interval, termed a tritone, a listener cannot invoke proximity in making judgments about tone pairs. Will perceptions of relative height, then, be ambiguous, or will some other principle be used to avoid ambiguity?

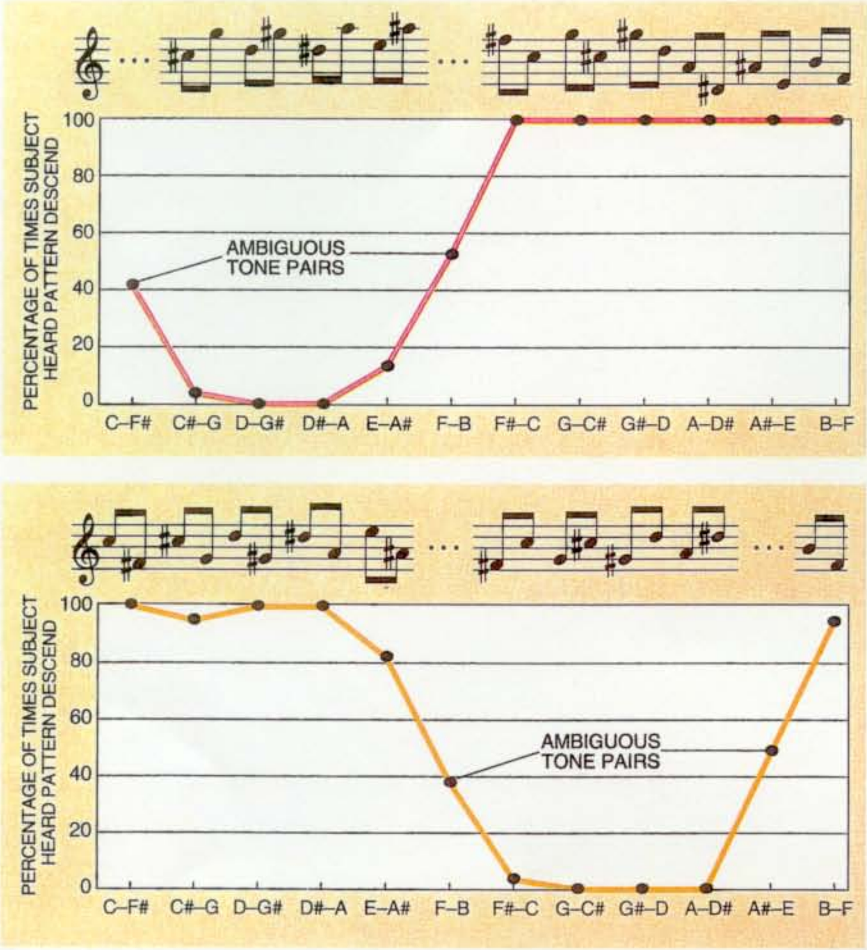
When I considered this question, it occurred to me that another cue was available to the perceptual system. A listener could establish absolute positions for the tones along the pitch class circle. For instance, we can envision the pitch classes to be the numbers on a clock face. Subjects might orient this clock face so that C is in the 12 o'clock position, C# in the one o'clock position and so on. They would hear C-F# (and

B-F and C#-G) as descending and F#-C (and F-B and G-C#) as ascending.

To examine this hypothesis, I presented subjects with just such tone pairs, which were generated by a computer program developed by F. Richard Moore of the University of California at San Diego. Each tone consisted of six sinusoidal components, all in the same pitch class. Subjects judged whether the tones ascended or descended. Their impressions were plotted as a function of pitch class of the first tone of the pair; hence, the pair C-F# would fall under pitch class C. The results strikingly confirmed the hypothesis. The judgments of most subjects showed orderly relations to the pitch class of the first tone: tones in one region of the circle were heard as higher than those in the opposite region.

A wholly unexpected finding emerged as well. The orientation of the pitch class circle with respect to pitch height varied radically from one subject to another. Some subjects heard, for example, the tone pair D-G# as descending, indicating that they oriented D in the upper half of the pitch circle (between the nine and three o'clock positions) and G# in the lower half. Others, however, heard the pattern ascend; they oriented D in the lower half of the circle and G# in the upper half. Subjects heard the illusion reverse itself when the pattern was transposed along the semitone scale.

These findings demonstrate that the pitch class circle is not flat with respect to pitch height. Rather, when listeners decide whether pairs of tones related by half an octave form ascending or descending patterns, their judgments are systematically related to the positions of the tones along the pitch class circle. There is, however, striking disagreement as to which region of the circle is tagged as the higher half and which the lower one. These findings, taken together, constitute the tritone paradox.



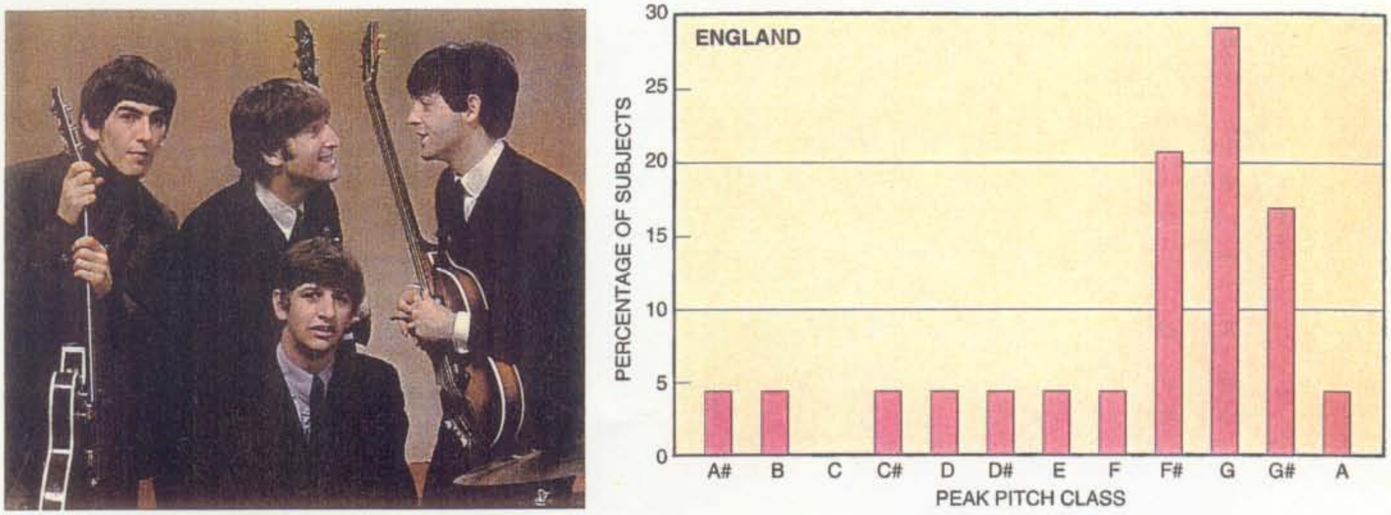
PERCEPTIONS of the tritone paradox by two subjects differed considerably. The first subject (*top*) clearly heard C#-G, D-G#, D#-A and E-A# as ascending but F#-C, G-C#, G#-D, A-D#, A#-E and B-F as descending. The second one (*bottom*) heard virtually the opposite: pairs heard by the first subject as ascending were perceived to descend, and vice versa. Both found some pairs ambiguous; that is, the tone pairs were heard to ascend and descend about equally often.

With William L. Kuyper of U.C.S.D. and Yuval Fisher of Cornell University, I performed a large-scale study of the tritone paradox. We selected a group of subjects who were U.C.S.D. undergraduates, had normal hearing and could reliably judge whether pairs of tones formed ascending or descending patterns. We found that the positions of the tones along the pitch class circle strongly influenced individual judgments. The direction of this tendency varied considerably from one subject to another, as expected. Furthermore, computer simulations showed that such perceptions exist to a highly significant extent in the general population.

I conducted other studies to see whether the phenomenon occurred when more complex patterns of tones were used. Specifically, I created the semitone paradox. The pattern consisted of two pairs of tones presented simultaneously; one pair ascended by a semitone, whereas the other pair descended by that interval. The tone pairs were diametrically opposed along the pitch class circle, so that again proximity could not be invoked in making judgments of relative height. Listeners generally perceived this pattern as two stepwise sequences that moved in opposite directions. Some listeners, however, heard the higher line of tones as ascending and the lower line as descending. Others perceived the reverse pattern.

Just as with the tritone paradox, judgments of the semitone paradox reflected an orderly relation between the perceived heights of the tones and their positions along the pitch class circle. In addition, the form of this relation varied substantially across subjects. For example, one subject heard the tones F, F#, G, G#, A and A# as higher than C, C#, D and D#. Yet for a second subject, the tones C#, D and D# were higher than the tones F, F#, G, G#, A, A# and B.

With my colleagues Moore and Mark B. Dolson, I also generated a more complex type of paradox by presenting three simultaneous pairs of tones to our subjects. For example, we played the tones D and B, followed by E and A and then F and G. Listeners generally perceived this pattern as two simultaneous melodies, one higher and one lower, that moved in opposite directions. Some listeners heard the higher melody as descending and the lower melody as ascending; others heard the opposite. When we transposed the pattern by half an octave, the subjects reported that



BRITISH AND CALIFORNIAN POPULATIONS perceive the tritone paradox in virtually opposite ways. Each graph shows the proportion of subjects that perceived a tone to reside at the peak region of the pitch class circle (that is, at the 12

the higher and lower melodies appeared to exchange positions. This melodic paradox shows that the subjects were perceptually preserving the relative heights of the different pitch classes.

I later presented the melodic paradox in six different keys: C, D, E, F#, G# and A#. These keys correspond to six equal steps along the pitch class circle. The subjects' perceptions of ascending and descending melodies depended on the key in which the tones were played. In another demonstration, I played the pattern as the keys shifted up in whole tone steps, so that the pattern was first in the key of C, then D and so on. Most people first heard the pattern one way; as the keys shifted, the pattern turned upside down and finally righted itself. In other words, the pattern appeared to rotate, in a fashion analogous to the rotation of shapes in vision.

These paradoxes lead to some fascinating conclusions about the auditory perceptual system. Clearly, the principle of equivalence under transposition, which investigators had thought to be universal, can be violated. Shifts in key most definitely affect the perception of certain patterns of tones. Another surprising conclusion concerns the phenomenon of absolute pitch, which is the ability to name a note just from hearing it. Musicians prize this presumed-to-be-rare faculty. Yet the experiments show that the ability is in greater supply than has been thought (at least in a partial form): individuals can perceive notes as higher or lower simply on the basis of pitch class.

The findings support research conducted in the early 1980s by Ernst Terhardt and Manfred Seewann of the Technical University of Munich and W. Dixon Ward of the University of Minnesota. These investigators found that musicians could generally determine whether well-known passages were played in the correct key, even though most of them did not have absolute pitch as conventionally defined. In fact, many succeeded even when the difference was as small as a semitone.

Finally, the paradoxes might have some implications for everyday listening to music. I found that different types of tone complexes—similar to groups of tones produced by natural instruments playing in octave relation—create the same paradoxes as do the more simple tones considered here. Furthermore, the effects endured when the tones were subjected to such time-varying manipulations as rapid fluctuations in pitch (vibratos), quick changes in loudness or fast decays.

Given the variety of sounds that produce these paradoxes, it is likely they also occur in music played by natural instruments. They could lead to subtle perceptual differences that could be of aesthetic importance in some pieces, especially orchestral works in which the composer has created an impression of ambiguity. Such ambiguities appear in Debussy's *Nocturnes.*

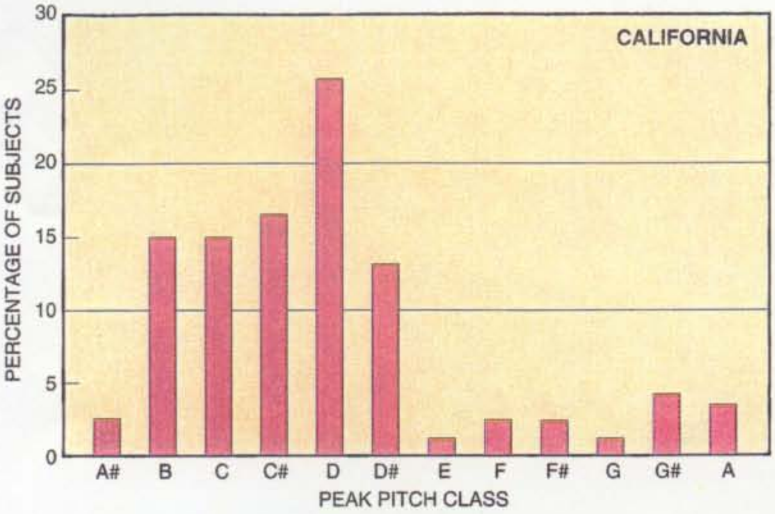
What can be the basis of this unexpected relation between pitch class and perceived height and of the individual differences in the manifestation of these paradoxes? The experiments conducted on the general population, where no correlation with musical training was found, indicate that the paradoxes are not musical in origin. In other studies, I ruled out simple characteristics of the hearing mechanism. For instance, I conducted experiments in which the odd-numbered components of each tone were played to one ear and the even-numbered ones to the other. The tritone paradox persisted, indicating that the brain produces the paradox by integrating information from both ears.

Based on these conclusions, and several informal observations, I conjectured that speech patterns might be responsible for the results. I had noticed that people from other countries often heard the paradox differently than did Californians. For example, individuals from the south of England tended to hear the tritone paradox in an opposite manner to that typical of subjects born in California: it appeared that what one group heard as ascending, the other heard as descending.

What might account for this observation? I hypothesized that people might have a long-term representation of the pitch range of their speaking voices. Included in this representation might be a preference for the octave band that contains the largest proportion of pitch values in their speech. I further hypothesized that the listener fixes the pitch classes defining this octave band at the highest positions along the pitch class circle (that is, near the 12 o'clock position). This definition determines the way the listener hears the paradoxes.

To examine this hypothesis, I undertook a study together with my colleagues Tom North and Lee Ray. We selected subjects whose judgments of the tritone paradox showed clear relations between pitch class and perceived height. We then recorded 15 minutes of spontaneous speech from each subject. From this recording, we determined the octave band containing the the largest number of pitch values in his or her speech.

Comparing the results from each subject, we found a significant correspondence between the pitch classes



o'clock position). In general, when one group heard a pattern descend, the other heard it ascend. It remains to be seen whether such differences have implications for the appreciation or performance of music.

defining this octave band for speech and those defining the highest position along the pitch class circle, as determined by each individual's judgments of the tritone paradox. The findings from this experiment substantiate the hypothesis that perception of the tritone paradox is based on the listener's representation of the pitch class circle—that is, on a kind of perceptual template. The orientation of this template is related to the pitch range of one's speaking voice.

How do these preferences arise? One interpretation would suggest that the listener's vocal range is completely innate. Indeed, some auditory illusions appear to result from differences at a basic neurological level. Research has shown, for instance, that the perception of some illusions correlates with the handedness of the listener [see "Musical Illusions," by Diana Deutsch; SCIENTIFIC AMERICAN, October 1975]. Another possibility is that a speech template could be acquired developmentally, through exposure to speech produced by others. Individuals would use this template to constrain their own speech output and to evaluate perceived speech.

The characteristics of such a template would therefore be expected to vary for those who speak in different languages or dialects, in a fashion similar to such other speech features as vowel quality. Following this line of reasoning, the orientation of the pitch class circle with respect to height should be similar for individuals who speak in the same language or dialect but should vary for those who speak in different languages or dialects.

The study of California undergraduates conducted by Kuyper, Fisher and me provided the initial evidence for the developmental acquisition of a speech template. Although no information was obtained concerning the linguistic backgrounds of these subjects, the majority had probably grown up in California and were from the same linguistic subculture. An orderly distribution of peak pitch class emerged among these subjects: C# and D occurred most frequently as peak pitch classes, followed by C and D#.

Given this preliminary finding, I carried out a more detailed study. I chose two groups for this purpose. The first consisted of 24 individuals who had grown up in California and spoke with the regional accent. The second was made up of 12 people from the south of England. In the experiment, I found that in the English group, F#, G and G# occurred most often as peak pitch classes. But in the California group, B, C, C#, D and D# occurred most frequently. Musical training appeared to have no effect, and neither did age or gender. This experiment provides strong support for the view that, through a learning process, the individual acquires a representation of the pitch class circle that has a particular orientation with respect to height.

It seems safe to assume that we employ such a template in our own speech and in the interpretation of the speech of others. Thus, a template based on pitch class rather than pitch height can be invoked by male and female speakers, even though they speak in different pitch ranges.

Some evolutionary value for such a template must exist. It could serve to provide a framework, common to a particular dialect, within which the pitch of a speaker's voice may be evaluated for emotion. A listener may also invoke the template in the communication of syntactic elements of speech.

It is likely that other disagreements in pitch perception await discovery. Unfortunately, we do not describe our musical experiences in terms that are precise or accurate enough for such differences to become apparent from informal observation. Only in the laboratory can researchers develop a clear idea of what a listener actually perceives and further define relations between speech and music. In many ways the studies confirm what some philosophers and musicians have argued for centuries: that one can achieve expressiveness in music by incorporating characteristics of the speaking voice, such as tempo and pitch range and variability. As the composer Mussorgsky wrote in his autobiography, "the function of art [is] the reproduction in musical sounds not merely of feelings, but first and foremost of human speech."

FURTHER READING

CIRCULARITY IN JUDGMENTS OF RELATIVE PITCH. Roger N. Shepard in *Journal of the Acoustical Society of America,* Vol. 36, No. 12, pages 2346–2353; December 1964.

AUDITORY PATTERN RECOGNITION. Diana Deutsch in *Handbook of Perception and Human Performance.* Edited by Kenneth R. Boff, Lloyd Kaufman and James P. Thomas. Wiley Interscience, 1986.

THE TRITONE PARADOX: AN INFLUENCE OF LANGUAGE ON MUSIC PERCEPTION. Diana Deutsch in *Music Perception,* Vol. 8, No. 4, pages 335–347; Summer 1991.

SOME NEW PITCH PARADOXES AND THEIR IMPLICATIONS. Diana Deutsch in *Philosophical Transactions of the Royal Society of London,* Series B, Vol. 336, No. 1278, pages 391–397; June 1992.