

10

THE PROCESSING OF PITCH COMBINATIONS

DIANA DEUTSCH

*Department of Psychology
University of California, San Diego
La Jolla, California*

I. INTRODUCTION

In this chapter, we shall examine ways in which pitch combinations are abstracted by the perceptual system. First, we shall inquire into the types of abstraction that give rise to the perception of local features, such as intervals, chords, and pitch classes. Such features can be considered analogous to those of orientation and angle size in vision. There has developed sufficient understanding of sensory physiology to justify speculation concerning how such abstractions are achieved by the nervous system. Other low-level abstractions result in the perception of global features, such as contour. Next, we shall consider how combinations of features are abstracted so as to give rise to perceptual equivalences and similarities. We shall then examine how these higher level abstractions are themselves combined according to various rules.

Investigations into mechanisms of visual shape perception have led to a distinction between early processes, in which many low-level abstractions are carried out in parallel, and later processes, in which questions are asked of these low-level abstractions based on hypotheses concerning the scene to be analyzed (Hanson & Riseman, 1978). The distinction between abstractions that are formed passively from “bottom up” and those that occur from “top down” is important in music also. As we shall see, both types of processes play important roles in musical shape analysis.

The final sections of the chapter are concerned with memory. We shall argue that musical patterns are retained in parallel in a number of memory systems, which correspond to different types and levels of abstraction, and that information from these different systems combine to determine memory judgments. We shall examine the interactions that take place within these systems, as well as how the outputs from these systems influence each other during retrieval.

II. FEATURE ABSTRACTION

A. OCTAVE EQUIVALENCE

It is clear that a strong perceptual similarity exists between tones that are separated by octaves; that is, those whose fundamental frequencies stand in the ratio of 2:1 (or a power of 2:1). Octave equivalence is implied in the musical systems of many different cultures (Nettl, 1956). In the Western musical scale, tones that stand in octave relation are given the same name, so that a tone is specified first by its position within the abstracted octave and then by the octave in which it occurs (G_3 , F_4 , and so on). In one version of Indian musical notation, a tone is also represented by a letter that designates its position within the octave, together with a dot or dots which designates the octave in which it occurs. So, for example, the symbols \dot{m} , m , \ddot{m} , \ddot{m} , and \ddot{m} represent tones that ascend by octave intervals. Furthermore, in Western tonal music, unisons and octaves are treated as harmonically interchangeable, and chord inversions are regarded as harmonically equivalent to their parent chords.

Various observations related to octave equivalence have been made. Baird (1917) and Bachem (1954) found that listeners with absolute pitch may sometimes place a note in the wrong octave, even though they name it correctly. Other researchers have demonstrated generalization of response to tones that stand in octave relation, in both people (Humphreys, 1939) including young infants (Demany & Armand, 1984) and animals (Blackwell & Schlosberg, 1943). In addition, interference effects in memory for pitch exhibit octave generalization (Deutsch, 1973b).

Given the perceptual similarity of tones that stand in octave relation, it has been suggested that musical pitch should be treated as a bidimensional attribute; the first dimension representing overall pitch level (*pitch height*) and the second dimension defining the position of the tone within the octave (*tone chroma* or *pitch class*) (Babbitt, 1960, 1965; Bachem, 1948; Deutsch, 1969, 1986a, 1992; Forte, 1973; M. Meyer, 1904, 1914; Révész, 1913; Risset, 1969, 1971; Ruckmick, 1929; Shepard, 1964, 1982). This suggestion is considered in detail in later sections of the chapter.

B. PERCEPTUAL EQUIVALENCE OF INTERVALS AND CHORDS

When two tones are presented either simultaneously or in succession, there results the perception of a musical interval, and intervals are perceived as being the same in size when the fundamental frequencies of their components stand in the same ratio. This principle forms an important basis of the traditional musical scale. The smallest unit of this scale is the semitone, which corresponds to a frequency ratio of approximately 1:1.06. Tone pairs that are separated by the same number of semitones are given the same name, such as major third, minor sixth, and so on.

Chords consisting of three or more tones are also classified in part by the ratios formed by their components. However, a simple listing of these ratios is not sufficient to define a chord. For instance, major and minor triads are perceptually quite distinct from each other, yet they are both composed of a major third (five semitones) a minor third (four semitones), and a perfect fifth (seven semitones). It is therefore of perceptual importance that the minor third lies above the major third in the major triad, and below it in the minor triad; this needs to be taken into account in considering how chords might be abstracted by the nervous system.

Given the principles of octave and interval equivalence, one might hypothesize that the perceptual equivalence of intervals would persist if their components were placed in different octaves. This assumption is frequently made by contemporary music theorists, who describe such intervals as in the same *interval class*. However, traditional music theory assumes that such equivalence holds for simultaneous but not successive intervals. Simultaneous intervals whose components have reversed their positions along the height dimension are treated as harmonically equivalent (Piston, 1948). Thus a simultaneous interval of n semitones is considered in a sense perceptually equivalent to one of $12 - n$ semitones.

Experimental evidence has been obtained for the perceptual similarity of inverted intervals. Plomp, Wagenaar, and Mimpfen (1973) had subjects identify intervals formed by simultaneous tone pairs and found that confusions occurred between those that were related by inversion (see also Deutsch & Roll, 1974). For the case of intervals formed by successive rather than simultaneous tone pairs, it appears that the perception of interval class does not occur directly, but rather through a process of hypothesis confirmation, in which the features that are directly apprehended are pitch class and interval. The evidence for this view is described later.

C. PROPOSED PHYSIOLOGICAL SUBSTRATES

Various models have been advanced to explain the perceptual equivalence of intervals and chords. Pitts and McCulloch (1947) hypothesized that the auditory cortex is composed of layers, with each layer containing a tonotopic projection of frequency-specific units. In each projection, units responding to frequencies that are related by equal intervals are spaced equal distances apart. These layers are arranged so as to produce columns of units that respond to the same frequencies. The authors further hypothesized that a set of fibers traverse this columnar mass parallel to each other in a slantwise direction; three such slanting fibers then define a three-note chord. This proposed mechanism could mediate the transposition of simultaneous intervals and chords, but would be unable to mediate transposition of successive intervals, nor could it account for the perceptual similarity of intervals and chords that are related by inversion.

An alternative hypothesis, suggested by Boomsalter and Creel (1961), was based on the volley theory of pitch perception (Wever & Bray, 1930). The authors pointed out that when two frequency combinations stand in the same ratio, they

should generate the same temporal pattern of firing. One temporal pattern would be produced by frequencies that stand in the ratio of 2:3, another in the ratio of 4:5, and so on. They proposed that the perceptual equivalence of simultaneous intervals is mediated by the recognition of such temporal patterns. One problem for this model is that it requires the nervous system to follow frequencies at much higher rates than have been demonstrated to occur (Edelman, Gall, & Cowan, 1988). As a further problem, the model cannot account for the perceptual equivalence of successive intervals and chords, nor for the perceptual similarity of intervals and chords that are related by inversion.

Deutsch (1969) proposed a neural network that would accomplish the abstraction of low-level pitch relationships so as to produce a number of equivalences that are found in the perception of music. This model was based on findings concerning the abstraction of low-level features by the visual system, such as orientation and angle size (Hubel & Wiesel, 1962).

The hypothesized neural network consists of two parallel channels, along each of which information is abstracted in two stages. An outline of this model is shown in Figure 1. The first channel mediates the perceptual equivalence of intervals and chords under transposition. In the first stage of abstraction along this channel, first-order units that respond to tones of specific pitch project in groups of two or three onto second-order units, which in consequence respond to specific intervals and chords (such as C_4 , E_4 , and G_4 , or D_5 and G_5). It is assumed that such linkages occur only between units underlying pitches that are separated by an octave or less. In the second stage of abstraction along this channel, second-order units project to third-order units in such a way that those second-order units that are activated by tones standing in the same relationship project onto the same unit. So, for example, all units that are activated by an ascending interval of four semitones (a major third) project onto one unit, all those that are activated by a descending interval of seven semitones (a perfect fifth) project onto a different unit, all those

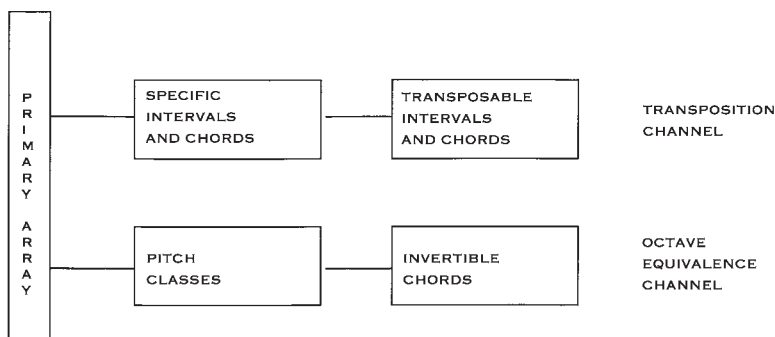


FIGURE 1 Model for the abstraction of pitch relationships. Pitch information is abstracted along two parallel channels; one mediating transposition and the other mediating octave equivalence. (Adapted from Deutsch, 1969. ©1969 by the American Psychological Association. Adapted with permission.)

that are activated by a major triad project onto yet a different unit, and so on (Figure 2).

The second channel mediates the perceptual equivalence of tones that stand in octave relation. In the first stage of abstraction along this channel, first-order units that respond to tones of specific pitch project to higher order units in such a way that those standing in octave relation project onto the same unit. These second-order units then respond to tones in a given pitch class, regardless of the octave in which they occur. In the second stage of abstraction along this channel, second-order units project in groups of two or three onto third-order units, which in consequence respond to combinations of pitch classes. Such units therefore mediate the perceptual similarity of intervals and chords that are related by inversion (Figure 3). This level of convergence is assumed to occur only for units that respond to simultaneously presented tones.

Although no attempt has been made to confirm this model at the neurophysiological level, the type of architecture that it postulates has been shown to exist in the auditory system. Neurons have been found that act as AND gates, as hypothesized for the transposition channel, and others as OR gates, as hypothesized for the octave equivalence channel. As another characteristic, auditory analyses are often carried out in parallel subsystems, each of which is organized in hierarchical fashion (see, e.g., Knudsen, du Lac, & Esterly, 1987; Schreiner, 1992; Suga, 1990; Sutter & Schreiner, 1991).

More recently, Bharucha (1987, 1991) has hypothesized a more elaborate neural network, whose basic architecture has features that are similar to those proposed by Deutsch (1969), and that also develops feature detectors for chords and keys as a result of passive exposure to the music of our tradition. This network is described in Chapter 11.

D. CONTOUR

We use global as well as specific cues in recognizing music. Such cues include, for example, overall pitch range, the distribution of interval sizes, and the relative proportions of ascending and descending intervals. Melodic contour plays a particularly important role here. As shown in Figure 4, melodies can be represented by their distinctive contours, even when interval sizes are not preserved.

One line of experimentation involving contour was initiated by Werner (1925). He reported that melodies could be recognized when they were transformed onto

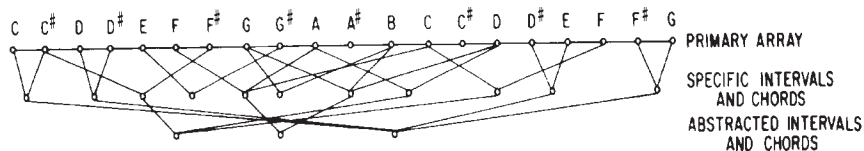


FIGURE 2 Two stages of abstraction along the transposition channel. (Adapted from Deutsch, 1969. ©1969 by the American Psychological Association. Adapted with permission.)

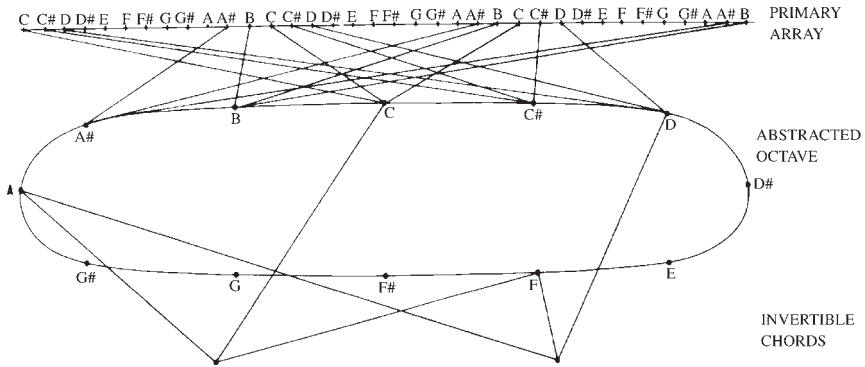


FIGURE 3 Two stages of abstraction along the octave-equivalence channel. (Adapted from Deutsch, 1969. ©1969 by the American Psychological Association. Adapted with permission.)

scales in which the octave was replaced by a different ratio, such as a fifth or two octaves, with these micro- or macro-octaves being divided into 12 equal intervals, so producing micro- or macro-scales. Later, Vicario (1983) carried out a study to determine how well listeners were able to recognize well-known melodies that had been transformed in this fashion. The results of this study are shown in Figure 5. As can be seen, although listeners were able to recognize such distorted melodies to some extent, the distortions did impair melody recognition, with the amount of impairment being a function of the degree of expansion or compression of the octave.

In another experiment, White (1960) found that listeners could recognize melodies to some extent when all the intervals were set to 1 semitone, so that only the sequence of directions of pitch change remained. Performance was enhanced when the relative sizes of the intervals were retained, but their absolute sizes were altered. Further studies have confirmed that contour can serve as a salient cue to melody recognition (see, e.g., Croonen, 1994; Dowling, 1978; Dowling & Fujitani, 1971; Edworthy, 1985; Idson & Massaro, 1978; and Kallman & Massaro, 1979).

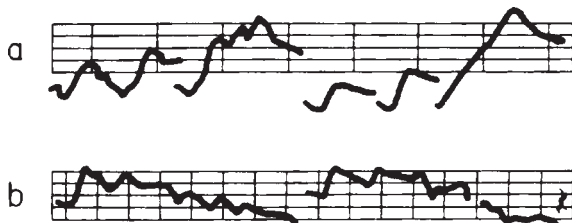


FIGURE 4 Contours from Beethoven piano sonatas as represented by Schoenberg: (a) from Sonata in C minor, Op. 10/I-III; (b) from Sonata in D, Op. 10/3-III, mm. 1-16. (From Schoenberg, 1967.)

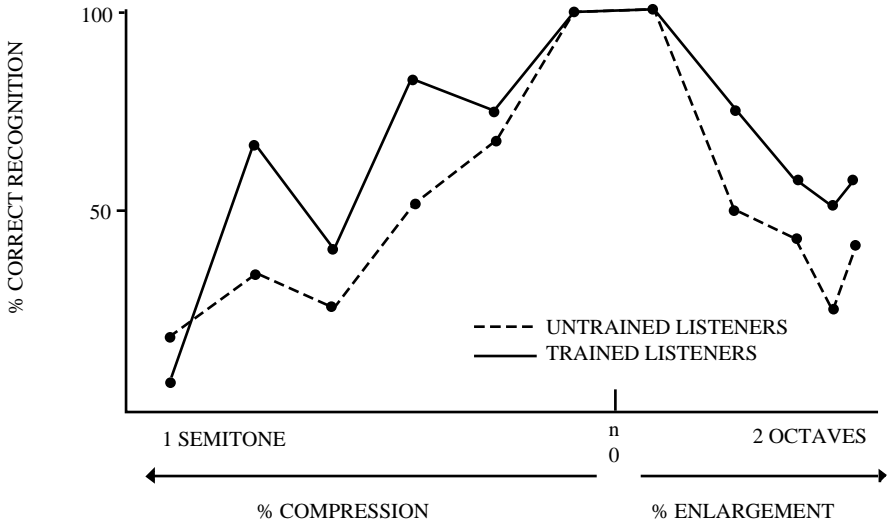


FIGURE 5 Percent correct recognition of melodies that have been transformed by compressing or enlarging the octave to differing extents. (Adapted from Vicario, 1983.)

E. MELODIC ARCHETYPES

L. B. Meyer (1973) has argued that there are a few schemata that, either through innate disposition or through experience with the music of our tradition, we search for and find in melodies. One such schema is characterized by a melodic leap, usually upward, which is followed by a gradual progression back to the first note. An example of such a “gap-fill” pattern is shown in the upper portion of Figure 6. Another such schema, which Meyer called a “changing-note pattern” consists of a stepwise movement around a prominent tone. For example, the pattern might begin on the tonic, move one step down, then two steps up and then one step down again (the sequence “do-ti-re-do”) as shown in the lower portion of Figure 6.

Meyer argued that such schemata enter into the hierarchical representation of melodies, particularly at higher levels, where they may be visually masked by the lower level patterns in the written score. He conjectured that listeners categorize melodies in accordance with the schemata that are involved at the highest structural level, regardless of the note-to-note structure of the melody.

As a test of this hypothesis, Rosner and Meyer (1982) took short melodic passages from recordings of classical music, and classified them as either gap-fill or changing-note at the highest structural level. Subjects without musical training learned to sort these melodies into two groups, even though the subjects were given no theoretical explanation for these two classifications. As the authors argued, this experiment indicated that we do indeed invoke such schemata in listening to music. This conclusion was reinforced in a later study by Rosner and Meyer

(a)

(b)

FIGURE 6 Illustrations of melodic archetypes. The upper illustration, from Geminiani's *Concerto Grosso* in E Minor, Op. 3, No. 3, shows a gap-fill pattern. The lower illustration, from Mozart's *Symphony No. 41* in C Major (K.551) shows a changing note pattern. See text for details. (Adapted from Rosner & Meyer, 1982.)

(1986). When subjects without formal musical training were asked to rate the degree of similarity between pairs of melodies that were either gap-fill or changing-note, they judged melodies with the same underlying schema as more similar to each other than those with different underlying schemata.

F. INTERVAL CLASS

When different two-tone combinations form the same interval by appropriate octave displacement, these combinations are held to be in the same interval class. For example, C_3 paired with D_5 , form the same interval class as G_2 paired with F_6 . The conditions under which interval class forms a basis for perceptual equivalence are complex ones. Experimental evidence for such equivalence has been obtained for simultaneous intervals, as described earlier (Deutsch & Roll, 1974; Plomp et al., 1973). Further evidence is provided by the general observation that we easily recognize root progressions of chords in different instantiations.

Where successive intervals are concerned, however, the issue is a complex one. If interval class were indeed a perceptual invariant, we should have no difficulty in recognizing a melody when its component tones are placed haphazardly in different octaves. This issue was examined in an experiment by Deutsch (1972c). The first half of the tune "Yankee Doodle" was generated in different versions. First, it was produced without transformation in each of three adjacent octaves. Second, it was generated in such a way that each tone was in its correct position within the octave (i.e., the interval classes were preserved) but the octave placement of the tones varied haphazardly within the same three octaves. Third, the tune was generated as a series of clicks, so that the pitch information was removed entirely but the rhythm remained. (A "scrambled-octaves" version of "Yankee Doodle" appears on the compact disc by Deutsch, 1995.)

The different versions of the tune were played to separate groups of subjects, who were given no clues as to its identity other than being assured that it was well known. Although the untransformed melody was universally recognized, the scrambled-octaves version was recognized no better than the version in which the pitch information was removed entirely. However, when the subjects were later told the name of the tune, and so knew what to listen for, they were able to follow the scrambled-octaves version to a large extent. This shows that the subjects were able to use pitch class to *confirm* the identity of the tune, although they had been unable to *recognize* it in the absence of cues on which to base a hypothesis.

This experiment provides strong evidence that perception of interval class, where successions of tones are concerned, requires the involvement of an active, "top-down" process, in which the listener matches each tone as it arrives with his or her image of the expected tone. On this line of reasoning, the extent to which interval class is perceived by the listener depends critically on his or her knowledge and expectations.

Other experimental findings have further supported the claim that interval class is not directly apprehended where successions of tones are concerned. In an experiment on short-term recognition of melodies, Deutsch (1979) presented listeners with a standard six-tone melody, followed by a comparison melody. The comparison melody was always transposed up 4 semitones from the standard. On half the trials the transposition was exact, and on the other half two of the tones in the transposed melody were permuted, while the melodic contour was unchanged.

There were four conditions in the experiment. In the first, the standard melody was played once, followed by the comparison melody. In the second, the standard melody was repeated six times before presentation of the comparison melody. In the third condition, the standard melody was again repeated six times, but now on half of the repetitions it was transposed intact an octave higher, and on the other half it was transposed intact an octave lower, so that the intervals forming the melody were preserved. In the fourth condition, the standard melody was again repeated six times, but now on each repetition the individual tones were placed alternately in the higher and lower octaves, so that the interval classes were preserved, but the intervals themselves were altered.

It was found that exact repetition of the standard melody resulted in a substantial improvement in recognition performance, and an improvement also occurred when the standard melody was repeated intact in the higher and lower octaves. However, when the standard melody was repeated in such a way that its tones alternated between the higher and lower octaves, performance was significantly poorer than when it was not repeated at all. This experiment provides further evidence that interval class cannot be considered a first-order perceptual feature. Repeating a set of intervals resulted in memory consolidation; however, repeating a set of interval classes did not do so.

The issue was further addressed in a study by Deutsch and Boulanger (1984). Musically trained subjects listened to novel melodic patterns which they recalled in musical notation. Examples of these patterns are given in Figure 7. Each pattern


Condition		Correct Notations
Higher octave		62.7%
Lower octave		67.5%
Both octaves		31.8%

FIGURE 7 Examples of sequences used in different conditions of the experiment on the effect of octave jumps on recall of melodic patterns. At the right are shown the percentages of tones that were correctly recalled in the correct serial positions in the different conditions. (Adapted from Deutsch & Boulanger, 1984. ©1984 by the Regents of the University of California.)

consisted of a haphazard ordering of the first six notes of the C-major scale. In the first condition, all the tones were taken from a higher octave; in the second, they were all taken from a lower octave. In the third condition, the individual tones alternated between these two octaves, so that roughly two thirds of the intervals formed by successive tones spanned more than an octave. The percentages of tones that were correctly notated in the correct serial positions in these different conditions are also shown in Figure 7. It can be seen that performance in the third condition was substantially poorer than in the other two.

The findings from these three experiments are in accordance with the two-channel model of Deutsch (1969), which assumes that neural linkages that underlie the abstraction of successive intervals occur only between units responding to pitches that are separated by no more than an octave. It is interesting in this regard to consider the use of octave jumps in traditional music. On the present line of reasoning, such jumps can be made with impunity, provided the musical setting is such that the displaced tone is anticipated by the listener. We should therefore expect that octave jumps would tend to be limited to such situations. Indeed, this appears to be the case. For example, a melodic line may be presented several times without transformation. A clear set of expectations having been established, a jump to a different octave occurs. The passage in Figure 8a, for instance, occurs after it has been presented several times without octave jumps. Another such situation is where the harmonic structure is clear and unambiguous, so that again the displaced tones are highly probable. This is illustrated in the segment in Figure 8b.

The technique of 12-tone composition uses very frequent octave jumps, and this raises the question of whether the listener does indeed identify as equivalent different presentations of the same tone row under octave displacement. Given the evidence and arguments outlined earlier, such recognition should be possible in principle, but only if the listener is very familiar with the material, or if its structure is such as to give rise to strong expectations (see also L. B. Meyer, 1973). Other work on melodic expectancy (see, in particular, L. B. Meyer, 1973, Narmour, 1990) is described elsewhere in this volume (see Chapter 12).



FIGURE 8 Two examples of octave jumps in traditional Western music. Here the jumps are readily processed. (a) From Beethoven, Rondo in C, Op. 5, No. 1; (b) from Beethoven, Sonata in C minor, Op. 10, No. 1.

III. HIGHER ORDER ABSTRACTIONS

Given that linkages are formed between first-order pitch elements, we may next inquire into how higher order abstractions are further derived so as to lead to perceptual equivalences and similarities. We recognize visual shapes when these differ in size, position in the visual field, and to some extent in orientation. What transformations result in analogous equivalences in music?

Theorists have long drawn analogies between perception of pitch relationships and relationships in visual space (Helmholtz, 1859/1954; Koffka, 1935; Mach, 1906). In contrast with visual space, however, pitch was conceived as represented along one dimension only. As Mach (1906/1959) wrote:

A tonal series is something which is an analogue of space, but is a space of one dimension limited in both directions and exhibiting no symmetry like that, for instance of a straight line running from right to left in a direction perpendicular to the median plane. It more resembles a vertical right line...

More recently, several investigators have shown that auditory analogues of visual grouping phenomena may be created by mapping one dimension of visual space into log frequency and the other into time (Bregman, 1978; Deutsch, 1975c; Divenyi & Hirsh, 1978; Van Noorden, 1975). The principle of proximity emerges clearly, for example, in the visual representation of the sequence shown in Figure 5 of Chapter 9. We may therefore inquire whether analogues of visual perceptual equivalences also exist in the way we represent music (Julesz & Hirsh, 1972).

A. TRANSPOSITION

Von Ehrenfels (1890), in his influential paper on form perception, pointed out that when a melody is transposed it retains its essential form, the *Gestaltqualität*, provided the relations among the individual tones are preserved. In this respect, he argued, melodies are similar to visual shapes; these retain their perceptual identities when they are translated to different locations in the visual field (Deese & Grindley, 1947).

A number of factors have been found to influence the extent to which a transposed and slightly altered melody is judged as similar to the original one. For example, when the original and transposed melodies can be interpreted as in the same key, and the successive tones comprising the melodies form the same number of steps along the diatonic scale, they are generally judged as very similar to each other; this is true whether or not the intervals forming the melody are preserved (Bartlett & Dowling, 1980; Dewitt & Crowder, 1986; Dowling, 1978, 1986; Takeuchi & Hulse, 1992; Van Egmond & Povel, 1994a). This can be taken to reflect the projection of pitch information onto overlearned alphabets, as described later (see Figures 12 and 13).

Several investigators have hypothesized that the extent to which a transposed melody is perceived as similar to the original is influenced by the key distance between them; for example, that a melody first played in C major will be judged as more similar when it is transposed to G major than to F# major (see, e.g., Cuddy & Cohen, 1976; Cuddy, Cohen, & Mewhort, 1981; Trainor & Trehub, 1993; Van Egmond, Povel, & Maris, in press). However, the explanations for apparent key distance effects are controversial (see, e.g., Takeuchi, 1994; Takeuchi & Hulse, 1992; Van Egmond & Povel, 1994a). Another factor that has been implicated is pitch distance. Several researchers have reported that the closer two melodies are in pitch range, the greater their perceived similarity (Francès, 1958/1988; Hershman, 1994; Van Egmond & Povel, 1994a, Van Egmond et al., in press). Finally the coding model of Deutsch and Feroe (1981) has been used as a predictor of perceived similarity between transposed melodies (Van Egmond & Povel, 1996); this work is described later.

B. INVERSION AND RETROGRESSION

We may next inquire whether further equivalences can be demonstrated for musical shapes that are analogous to their visuospatial counterparts. Schoenberg (1951) argued that transformations similar to rotation and reflection in vision result in perceptual equivalences in music also. He wrote:

The unity of musical space demands an absolute and unitary perception. In this space ... there is no absolute down, no right or left, forward or backward ... Just as our mind always recognizes, for instance, a knife, a bottle or a watch, regardless of its position, and can reproduce it in the imagination in every possible position, even so a musical creator's mind can operate subconsciously with a row of tones, regardless of their direction, regardless of the way in which a mirror might show the mutual relations, which remain a given quantity.

This statement may be compared with Helmholtz's (1844) description of imagined visuospatial transformations:

Equipped with an awareness of the physical form of an object, we can clearly imagine all the perspective images which we may expect upon viewing it from this or that side (see Warren & Warren, 1968, p. 252)

On this basis, Schoenberg proposed that a row of tones may be recognized as equivalent when it is transformed in such a way that all ascending intervals be-

come descending ones and vice versa (“inversion”), when it is presented in reverse order (“retrogression”), or when it is transformed by both these operations (“retrograde-inversion”). Figure 9 illustrates Schoenberg’s use of his theory in compositional practice. As Schoenberg (1951) wrote:

The employment of these mirror forms corresponds to the principle of the absolute and unitary perception of musical space.

Schoenberg did not conceive of the vertical dimension of musical space simply as pitch, but rather as pitch class. His assumptions of perceptual equivalence under transposition, retrogression, inversion, and octave displacement are fundamental to 12-tone composition (Babbitt, 1960, 1965). In this procedure, a given ordering of the 12 tones within the octave is adopted. The tone row is repeatedly presented throughout the piece; however, the above transformations are allowed on each

The figure consists of two main parts. The upper part is a diagram showing 12 tones arranged in a semi-circular arc. The tones are labeled with numbers 1 through 12 in the center, and 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 on the outer edge. The lower part is a musical score for piano, showing the 'Basic Set' and 'Retrograde Set' in the upper staves, and 'Inversion' and 'Retrograde Inversion' in the lower staves. The lower staves are connected to the top diagram by curved lines, illustrating the equivalence relations between the different forms of the tone row.

FIGURE 9 Schoenberg’s illustration of his theory of equivalence relations between pitch structures, taken from his *Wind Quartet*, Op. 26. (From Schoenberg, 1951.)

presentation. It is assumed that the row is perceived as an abstraction in its different manifestations.

Whether such transformations indeed result in perceptual equivalences is debatable. In the visual case, we must have evolved mechanisms that preserve the perceptual identities of objects regardless of their orientation relative to the observer. An analogous argument cannot be made for inversion and retrogression of sound patterns. A second doubt is based on general experience. Sound sequences may become unrecognizable when they are reversed in time, as the reader can confirm by attempting to decode a segment of speech when it is played backward. Furthermore, many inverted three-note combinations are perceptually very dissimilar to the combinations from which they are derived. For example, a minor triad is an inversion of a major triad, yet the two are perceptually quite distinct from each other. It would appear that when inverted and retrograde patterns are recognized, this is accomplished at a level of abstraction that is equivalent to the one that allows us to recite a segment of the alphabet backwards or to invert a series of numbers (Deutsch & Feroe, 1981). For other discussions of the perceptual status of 12-tone compositions, see Krumhansl, Sandell, and Sergeant (1987), Francès (1958/1988), and in particular Thomson (1991).

Other theorists have proposed representations of pitch relationships in terms of distances in multidimensional space. For example, in order to capture the close perceptual similarity between tones that stand in octave relation, it has been suggested that pitch be represented as a helix, with the vertical axis corresponding to pitch height and tones separated by octaves lying closest within each turn of the helix (Drobisch, 1855; Pickler, 1966; Shepard, 1964, 1982; Ueda & Ohgushi, 1987). The helical model is discussed in detail in Section IV.

More elaborate representations have also been proposed that would capture the complex patterns of pitch relationships that we invoke in listening to tonal music. For example, Longuet-Higgins (1962a, 1962b, 1978) has suggested that "tonal space" may be characterized as a three-dimensional array: Tones that are adjacent along the first dimension are separated by fifths, those adjacent along the second dimension by major thirds, and those adjacent along the third dimension by octaves. The intervals of tonal music then appear as vectors in this tonal space. If tones that are related by octaves are treated as equivalent, then an array such as that shown in Figure 10 is obtained. As can be observed, closely related tones, such as comprise the C-major scale, form a compact group in this array, so that a key can be defined as a neighborhood in tonal space. Similar representations have been proposed by others, such as Hall (1974) Balzano (1980), and Shepard (1982).

The spatial modeling of pitch relationships has a long tradition among music theorists. In particular, 18th century theorists developed circular configurations that would capture degrees of modulation between keys. In these models, adjacent positions along such circles depict close modulations, and positions that are further removed depict more distant ones. Later theorists such as Weber (1824) and Schoenberg (1954/1969) have produced related spatial models (Werts, 1983).

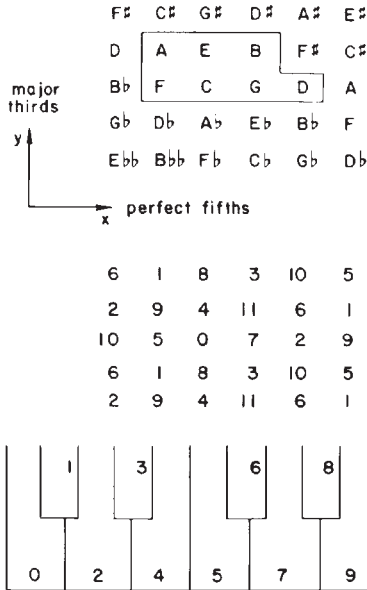


FIGURE 10 Array hypothesized by Longuet-Higgins for the representation of tonal space. (From Longuet-Higgins, 1978.)

C. PITCH ALPHABETS

An alternative model for the representation of pitch relationships in tonal music was proposed by Deutsch and Feroe (1981). The model assumes that, through extensive exposure to such music, the listener acquires a repertoire of hierarchically embedded alphabets, most prominently the chromatic scale, diatonic scales, and triads. An outline of the model is shown in Figure 11, taking as an example the key of C major. At the lowest level, the chromatic alphabet serves as the parent alphabet from which families of subalphabets are derived. The major and minor scales are represented at the next-higher level; these can be expressed in terms of proximal distances along the chromatic alphabet. Triads are represented at the next-higher level; these can be expressed in terms of proximal distances along diatonic alphabets. Lerdahl (1988) has proposed an elaboration of Deutsch and Feroe's hierarchy of alphabets that also takes account of a number of other characteristics of tonal music, such as patterns of proximity between chords.

Compositional practice reflects our use of such overlearned alphabets. For example, in the short-term transposition of motives, the number of steps along an alphabet is often preserved, and even when such transpositions result in alterations in interval size, they still appear appropriate to the listener. Figures 12 and 13 give two such examples. The first, from a Bach fugue, shows a motive that traverses the D-major scale four times in succession, each time beginning on a different posi-

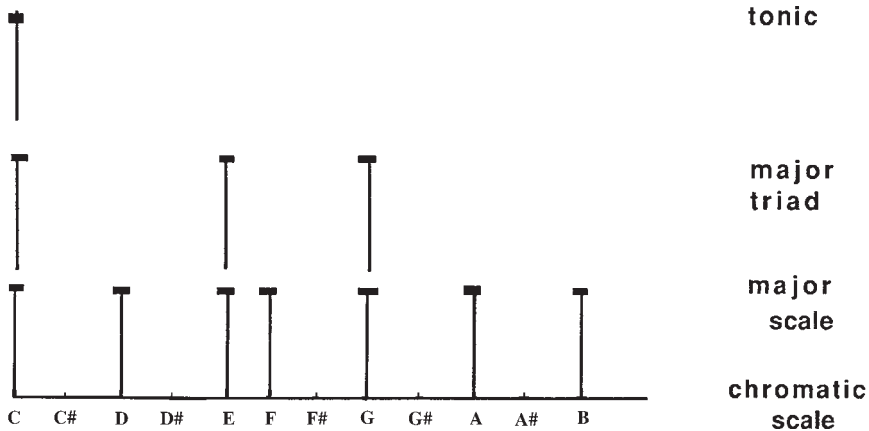


FIGURE 11 A hierarchy of embedded pitch alphabets. (Adapted from Deutsch & Feroe, 1981. ©1981 by the American Psychological Association. Adapted with permission.)

tion along this scale. The second, from a Schubert impromptu, shows a motive that traverses the A \flat -minor triad five times in succession, each time beginning at different positions along this triad.

The hypothesis that pitch structures are represented in terms of such embedded alphabets has received experimental support. Deutsch (1980) asked subjects to listen to sequences of tones drawn from such alphabets, and to recall what they heard in musical notation. It was found that when errors in notation occurred, they rarely departed from the alphabet that had been presented. In general, sequences were recalled very accurately when they could be simply represented as hierarchical structures, with different pitch alphabets at different levels of the hierarchy (see later).

Further evidence comes from reports that melodies were better remembered when they were composed only of tones in a particular diatonic set than when they also included tones outside the set (Cuddy et al., 1981; Dowling, 1991; Francès, 1958/1988). Presumably, adhering to a diatonic set made it likely that the listener would invoke a key, and so use overlearned pitch alphabets as an aid to memory. It has also been reported that changing the context of a melody so as to suggest a different key rendered the melody more difficult to recognize (Dowling, 1986). Yet other studies have found that transpositions of melodies that did not involve a change in key were judged as very similar to the original melodies, regardless of whether or not the intervals were preserved (Bartlett & Dowling, 1980; Dewitt & Crowder, 1986; Dowling, 1978, 1986; Takeuchi & Hulse, 1992; Van Egmond & Povel, 1994). In addition, an alteration in a melody was easier to detect when it could be interpreted as a departure from its key, and so involved a departure from the alphabets appropriate to the key (Francès, 1958/1988, Dewar, Cuddy, & Mewhort, 1977; Dowling, 1978).

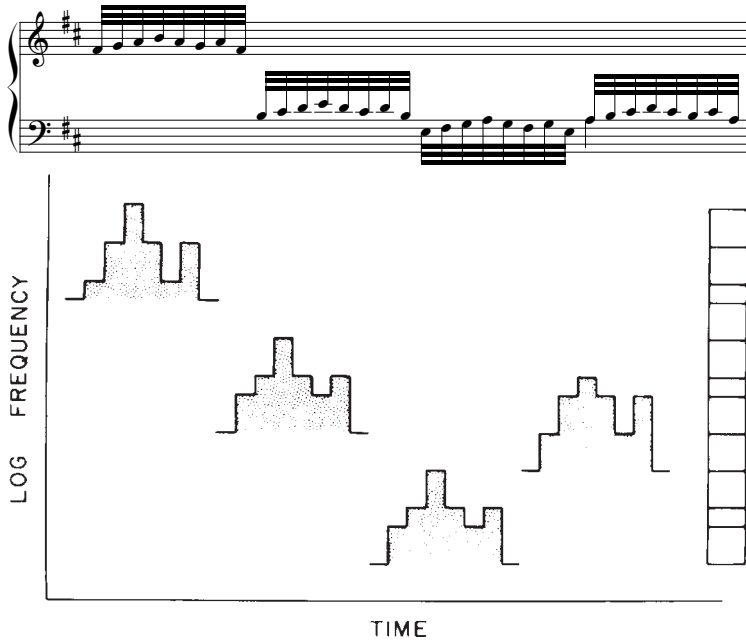


FIGURE 12 Transposition along the alphabet of the D-major scale. The same pattern is presented four times in succession at different positions along the scale. Because the major scale consists of unequal intervals, there result differences in the intervals comprising the pattern. The ladder at the right displays the scale. From J. S. Bach, *The Well-Tempered Clavier*, Book 1, Fugue V. (From Deutsch, 1977.)

An important point originally made by music theorists (see in particular, L. B. Meyer, 1956) is that in the context of an established key, the various scale degrees tend to be assigned differing levels of prominence by the listener. The tonic is generally perceived as the most prominent, followed by the mediant and dominant, and these in turn are followed by the remaining scale degrees, with the other tones of the chromatic scale lowest in the hierarchy.

Laboratory findings have provided evidence that such hierarchies of prominence influence our processing of musical patterns in various ways. They have been found to influence perceived relatedness between tones (Krumhansl, 1979, 1990a), patterns of confusion between tones in short-term memory (Krumhansl, 1979), expected continuations of melodies (Schmuckler, 1989), and the time taken to attribute keys to musical segments (Janata & Reisberg, 1988). Such hierarchies of prominence also correlate with the frequency of occurrence of the different scale degrees in the tonal music of our tradition (Krumhansl, 1990a).

Music theorists have also pointed out that there are hierarchies of prominence for triads within established keys, with the triad on the tonic heard as the most prominent, followed by the triad on the dominant, then the subdominant, and so on. In addition, it appears that the harmonic functions of chords influence the ex-

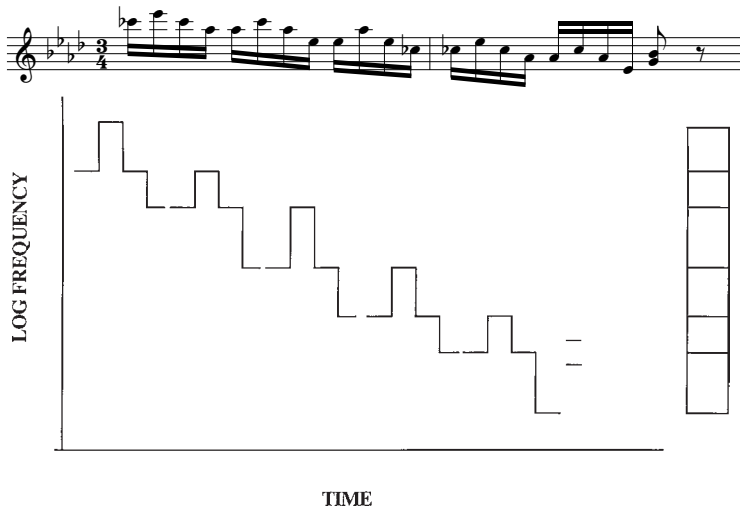


FIGURE 13 Transposition along the alphabet of the A^b -minor triad. The same pattern is presented five times in succession, at different positions along this triad. Because the triad consists of uneven intervals, there result differences in the intervals comprising the pattern. The ladder at the right displays the triad. From F. Schubert, *Four Impromptus*, Op. 90, No. IV.

tent to which they are perceived as related to each other (Krumhansl, Bharucha, & Kessler, 1982; Bharucha & Krumhansl, 1983) and the probability that they will be confused in memory (Bharucha & Krumhansl, 1983; Krumhansl & Castellano, 1983). Furthermore, alterations in the tonal context in which chords are presented affect patterns of memory confusion and judgments of relatedness between them (Bharucha & Krumhansl, 1983). Such hierarchies of prominence for chords within a key also appear to correlate with their frequency of occurrence in Western tonal music also (Krumhansl, 1990a).

D. HIERARCHICAL ENCODING

We now turn to the question of how pitch sequences in tonal music are encoded and retained in memory. Music theorists have argued that the tonal music of our tradition is composed of segments that are organized in hierarchical fashion (see, e.g., Lerdahl & Jackendoff, 1983; L. B. Meyer, 1956, 1973; Narmour, 1990; Schenker, 1956). It is reasonable to suppose that this form of organization reflects the ways in which musical information is encoded and retained in memory. As Greeno and Simon (1974) point out, we appear to retain many different types of information as hierarchies. In some instances, the information is stored in the form of concepts that refer to classes (Collins & Quillian, 1972). We also appear to retain hierarchies of rules (Gagné, 1962; Scandura, 1970), of programs (Miller, Galanter, & Pribram, 1960) and of goals in problem solving (Ernst & Newell, 1969). Visual scenes appear to be retained as hierarchies of subscenes (Hanson &

Riseman, 1978; Navon, 1977; Palmer, 1977; Winston, 1973). The phrase structure of a sentence lends itself readily to hierarchical interpretations (Miller & Chomsky, 1963; Yngve, 1960).

Restle (1970) and Restle and Brown (1970) have provided experimental evidence that we readily acquire serial patterns as hierarchies that reflect the structures of these patterns. Parallel theoretical developments by Simon and his colleagues (Simon, 1972; Simon & Kotovsky, 1963; Simon & Sumner, 1968) and by others (Leewenberg, 1971; Jones, 1974, 1978; Vitz & Todd, 1967, 1969) have addressed the ways in which we acquire and retain serial patterns in terms of hierarchies of operators.

Deutsch and Feroe (1981) have proposed a model of the way in which we represent pitch sequences in tonal music. In essence, the model can be characterized as a hierarchical network, at each level of which structural units are represented as organized sets of elements. Elements that are present at any given level are elaborated by further elements so as to form structural units at the next-lower level, until the lowest level is reached. The model also assumes that Gestalt principles of perceptual organization, such as proximity and good continuation, contribute to organization at each hierarchical level.

The model invokes hierarchies of operators, and a simplified version is as follows:

1. A structure is notated as $(A_1, A_2, \dots, A_{l-2}, A_{l-1}, *, A_{l+1}, A_{l+2}, \dots, A_n)$, where A_j is one of the operators n , p , s , n^i , or p^i . (A string of length k of an operator A is abbreviated kA .)

2. Each structure $(A_1, A_2, \dots, *, \dots, A_n)$ has associated with it an alphabet, α . The combination of a structure and an alphabet is called a *sequence* (or *subsequence*). This, together with the reference element r , produces a *sequence of notes*.

3. The effect of each operator in a structure is determined by that of the operator closest to it, but on the same side as the asterisk. Thus the operator n refers to traversing one step up the alphabet associated with the structure. The operator p refers to traversing one step down this alphabet. The operator s refers to remaining in the same position. The two operators n^i and p^i refer to traversing up or down i steps along the alphabet, respectively.

4. The values of the sequence of notes $(A_1, A_2, \dots, *, \dots, A_n)$, α , r , where α is the alphabet and r the reference element, are obtained by taking the value of the asterisk to be that of r .

5. To produce another sequence from the two sequences $A = (A_1, A_2, \dots, *, \dots, A_m)$, α , and $B = (B_1, B_2, \dots, *, \dots, B_n)$, β , where α and β are two alphabets, we define the compound operator pr (prime). $A[pr]B;r$, where r is the reference element, refers to assigning values to the notes produced from $(B_1, B_2, \dots, *, \dots, B_n)$ such that the value of $*$ is the same as the value of A_1 , when the sequence A is applied to the reference element r . Values are then assigned to the notes produced from $(B_1, B_2, \dots, *, \dots, B_n)$ such that the value of $*$ is the same as the value of A_2 , and so on. This gives a sequence of length $m \times n$. Other compound operators such as inv (inversion) and ret (retrograde) are analogously defined.

To characterize the advantages of the model, let us consider the sequence shown in Figure 14b. This may, in principle, be represented in terms of steps that traverse the chromatic scale: A basic subsequence that consists of one step up this scale is presented four times in succession; the second presentation being four steps up from the first, the third being three steps up from the second, and the fourth five steps up from the third. However, this analysis assigns prominence to the basic subsequence, and does not relate the various transpositions to each other in a musically meaningful way.

In contrast, a musical analysis of this pattern would describe it as on two hierarchical levels, in which a higher level subsequence is elaborated by a lower level subsequence. At the higher level, shown on Figure 14a, there is an arpeggiation that ascends through the C-major triad (the notes C-E-G-C). At the lower level, shown in Figure 14b, a neighbor embellishment precedes each note of the triad, so creating a two-note pattern. This hierarchical structure is represented in the dia-

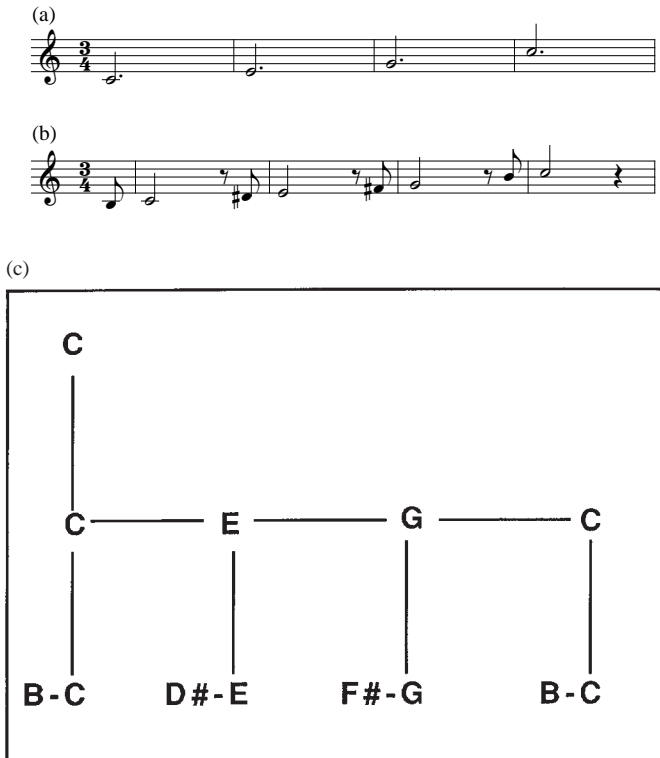


FIGURE 14 A series of pitches represented on two hierarchical levels. (a) At the higher level, there is an arpeggiation of the C-major triad. (b) At the lower level, each note of the triad is preceded by one a semitone lower, so forming a two-note pattern. (c) This hierarchical structure shown in diagram form. (Adapted from Deutsch & Ferøe, 1981. ©1981 by the American Psychological Association. Adapted with permission.)

gram in Figure 14c. According to the formalism just outlined, the pattern can be represented as:

$$\begin{aligned} A &= (*, 3n) C_{tr} \\ B &= (p, *) Cr \\ S &= A [pr]B, C_4 \end{aligned}$$

where C_{tr} represents the C-major triad, Cr the chromatic scale, and C_4 the reference element.

A more complex example is given in Figure 15 and is represented on three structural levels as follows:

$$\begin{aligned} A &= (*, 4p) b_{tr} \\ B &= (*, n, p) b_{tr} \\ S &= A [pr](B, 4(*))[inv, 5pr](B, (*)) D_5 \end{aligned}$$

In many other hierarchical representations of music, such as the one proposed by Schenker, and the other coded element models referred to earlier, elements at all but the lowest level are rule systems rather than actual notes. In contrast, in the present model, a sequence of notes (or subsequence) is realized at each structural level. This property confers a number of processing advantages to the listener. For example, it enables the encoding of such subsequences in terms of laws of figural goodness, such as proximity and good continuation, and also enables the invocation of melodic archetypes such as the gap-fill and changing-note patterns, described by L. B. Meyer (1973). In addition, it enables different pitch alphabets to be invoked at different structural levels, so facilitating the process of generating hierarchical representations from the musical patterns that we hear. As a related point, notes that are present at any given level are also present at all levels below it. In consequence, the higher the level at which a note is represented, the more often and so the more firmly it is represented. This has the consequence that higher level

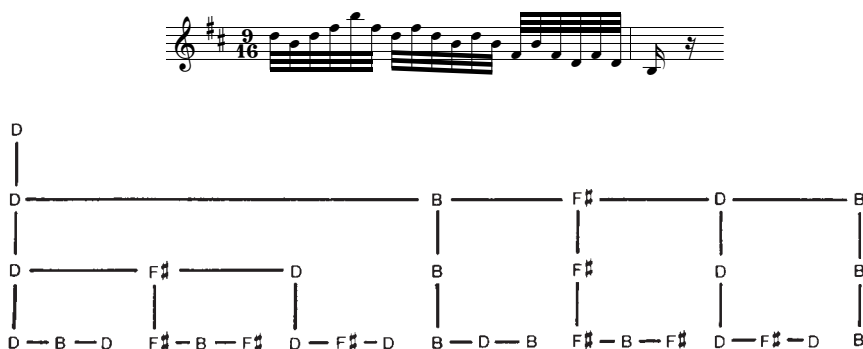


FIGURE 15 More complex hierarchy illustrating the Deutsch and Feroe model. Passage is from Bach's Sinfonia 15, BWV 801. (From Deutsch & Feroe, 1981. ©1981 by the American Psychological Association. Adapted with permission.)

subsequences act to cement lower level subsequences together. Another processing advantage of the model is that it enables sequences at different structural levels to be encoded as chunks of a few items each, which in turn is conducive to good memory performance (Estes, 1972; Wickelgren, 1967).

One hypothesis that arises from this model is that sequences of tones should be processed more easily when they can be parsimoniously represented in accordance with its rules. This hypothesis was examined in an experiment by Deutsch (1980). Musically trained listeners were presented with sequences of tones, which they recalled in musical notation. Examples of the sequences are shown in Figure 16. The passage in Figure 16a (a “structured sequence”) consists of a higher level subsequence of four elements that acts on a lower level subsequence of three elements. The passage in Figure 16b (an “unstructured sequence”) consists of a haphazard reordering of the one in Figure 16a, and does not lend itself to a similar parsimonious representation. It was predicted, on the basis of the model, that the structured sequences would be notated more accurately than the unstructured ones.

Another factor was also examined in this experiment. It has been found in studies using strings of verbal materials that we tend to recall such strings in accordance with their temporal grouping (Bower & Winzenz, 1969; McLean & Gregg, 1967; Mueller & Schumann, 1894). This effect was found to be so powerful as to offset grouping by meaning (Bower & Springston, 1970). Analogous results were also obtained using nonverbal materials (Dowling, 1973; Handel, 1973; Restle, 1972). It was predicted, therefore, that temporal grouping would affect ease of recall of the present tonal sequences also. In particular, temporal grouping in accordance with pitch structure was expected to enhance performance, whereas grouping in conflict with pitch structure was expected to result in performance decrements.

Given these considerations, sequences such as these were each presented in three temporal configurations (Figure 17). In the first, the tones were spaced at equal intervals; in the second, they were spaced in four groups of three, so that they were segmented in accordance with pitch structure; in the third, they were



FIGURE 16 Examples of sequences used in the experiment to study utilization of pitch structure in recall. Sequence (a) can be represented parsimoniously as a higher level subsequence of four elements that acts on a lower level subsequence of three elements. Sequence (b) consists of a haphazard reordering of the notes in sequence (a) and cannot be represented in such a parsimonious fashion. (Adapted from Deutsch, 1980.)



FIGURE 17 Types of temporal segmentation used in the experiment to study the utilization of pitch structure in recall. (a) Sequence unsegmented. (b) Sequence segmented in groups of three, so that segmentation is in accordance with pitch structure. (c) Sequence segmented in groups of four, so that segmentation is in conflict with pitch structure.

spaced in three groups of four, so that they were segmented in conflict with pitch structure.

Large effects of pitch structure and temporal segmentation were obtained. For structured sequences that were segmented in accordance with pitch structure, performance levels were very high. For structured sequences that were unsegmented, performance levels were still very high, though slightly lower. For structured sequences that were segmented in conflict with pitch structure, however, performance levels were much lower. For unstructured sequences, performance levels were considerably lower than for structured sequences that were segmented in accordance with their structure or that were unsegmented, but were in the same range as for structured sequences that were segmented in conflict with pitch structure.

Figure 18 shows the percentages of tones that were correctly recalled in their correct serial positions in the different conditions of the experiment. Typical bow-shaped curves are apparent, and in addition, discontinuities occur at the boundaries between temporal groupings. This pattern of results indicates that the subjects tended to encode the temporal groupings as chunks, which were retained or lost independently of each other. This pattern is very similar to that found by others with the use of verbal materials (Bower & Winzencz, 1969).

The transition shift probability (TSP) provides a further measure of interitem association. This is defined as the joint probability of either an error following a correct response on the previous item, or of a correct response following an error on the previous item (Bower & Springston, 1970). If groups of elements tend to be retained or lost as chunks, we should expect the TSP values to be smaller for transitions within a chunk, and larger for the transition into the first element of a chunk. It was indeed found that TSPs were larger on the first element of each temporal grouping than on other elements. This is as expected on the hypothesis that temporal groupings serve to define subjective chunks that are retained or lost independently of each other.

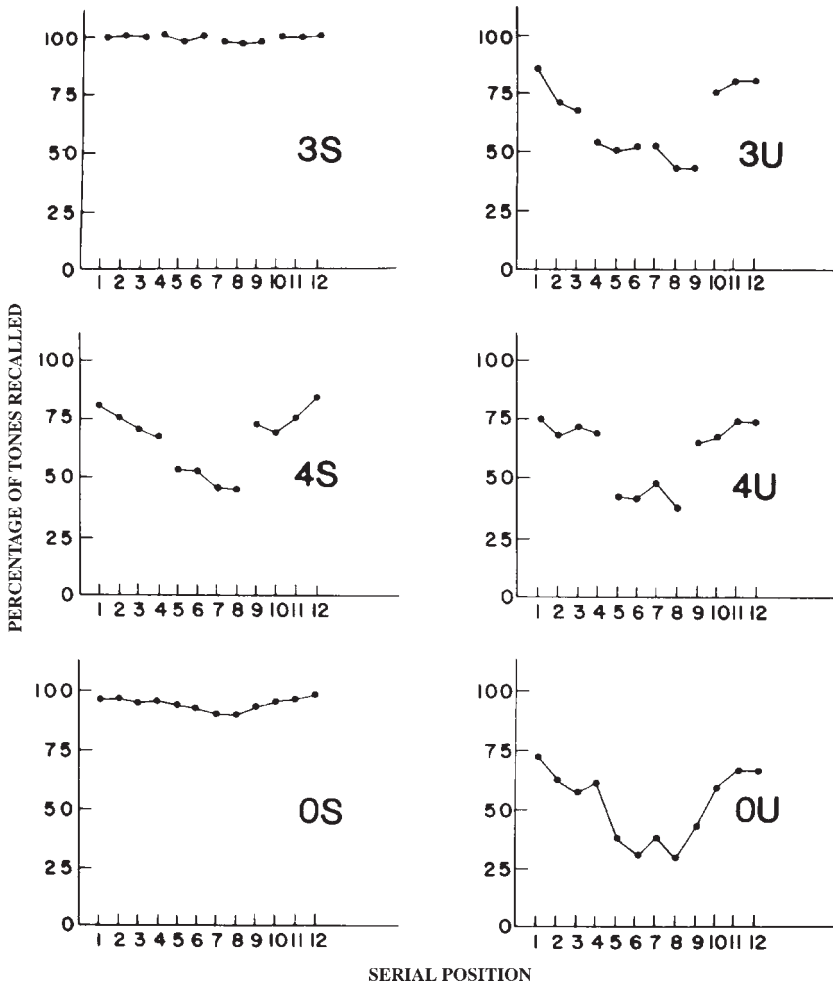


FIGURE 18 Serial position curves for the different conditions of the experiment to study the utilization of pitch structure in recall. 3: Temporal segmentation in groups of three. 4: Temporal segmentation in groups of four. O: No temporal segmentation. S: Structured sequence. U: Unstructured sequence. (From Deutsch, 1980.)

In general, these results provide strong evidence that listeners perceive hierarchical structures that are present in tonal sequences, and that they use such structures in recall. For the structured sequences used here, the listener needed only to retain two chunks of three or four items each; however, for the unstructured sequences, no such parsimonious encoding was possible. The error rates for the unstructured sequences were much higher than for the structured ones, in accordance with the hypothesis that they imposed a much heavier memory load.

Another study was carried out by Van Egmond and Povel (1996). A paired comparison paradigm was used to investigate perceived similarities between mel-

odies and their transpositions, when the latter had been altered in various ways. The Deutsch and Feroe model was used as a qualitative predictor of the degree of perceived similarity between the original and transposed melodies. The authors hypothesized that the larger the number of items by which the codes for the original and transposed melodies differed, the more dissimilar the two melodies would appear.

More specifically, Van Egmond and Povel predicted that an exact transposition would be judged as most similar to the original melody, because its code would differ only in terms of one item: the key. For a transposition that was chromatically altered, the prediction concerning perceived similarity would depend on whether the transposed melody could be represented parsimoniously in the same key as the original. If it could be so represented, then its code would differ in terms of only one item: the reference element. If it could not be so represented, then its code would differ in terms of two items: the key and the reference element. Finally, a transposition that was diatonically altered would be judged as most dissimilar to the original melody, because its code would differ in terms of six items: the key and five structure operators.

The experimental findings confirmed the hypothesis. Exact transpositions were judged as most similar to the original melodies. Chromatically altered transpositions that could be interpreted as in the same key as the original melodies were judged as more similar than were those that could not be so interpreted. Transpositions that were diatonically altered were judged as more dissimilar than were chromatically altered transpositions.

Two other studies are relevant here. Oura (1991) presented subjects with a melody, which they recalled in musical notation. She found that tones that were represented at higher structural levels were recalled better than those that were represented at lower levels. Dibben (1994) had subjects listen to a musical segment, and then to a pair of reductions, and she asked them to determine which reduction best matched the original. She found that the subjects chose the version that matched the full segment at higher structural levels. The findings from both these studies are in accordance with the prediction from Deutsch and Feroe (1981) that the higher up a note or sequence of notes is represented, the more often it is represented, and so the more firmly it should be embedded in memory.

So far we have been considering the processing of a single melodic line. However, tonal music generally involves several such lines, and even where only one is presented, a harmonic progression is generally implied. We can assume that such progressions are also encoded in hierarchical fashion. In addition, the use of parallel linear sequences, which must themselves combine to form an acceptable harmonic sequence, places constraints on the choice of elements in each sequence; this in turn serves to reduce the processing load.

E. KEY ATTRIBUTION

We now consider the process whereby the listener attributes a key to a passage. As is evident from general experience, we tend to make such decisions quickly

and accurately, often on the basis of very little information. However, there is much disagreement concerning which cues we use in making these decisions, and the relative weightings that we give to these cues.

One approach to this issue stresses the identities of the pitches in the passage, independent of their orderings. For example, it is assumed that when presented with the diatonic collection C, D, E, F, G, A, and B, the listener will attribute the key of C major. It is further hypothesized that when presented with other pitches taken from the chromatic scale (such as F \sharp or D \sharp), the listener will be more likely to reject the hypothesis of C major and will search instead for a different key (see, e.g., Longuet-Higgins, 1962a, 1962b, 1978).

This approach accords well with general experience; however, the experimental evidence indicates that other factors also play a role. In one experiment, West and Fryer (1990) presented subjects with a quasi-random ordering of the tones of a diatonic scale, followed by a probe tone. The subjects judged how suitable the probe tone was as a tonic, in the context of the sequence that had been presented. Even musically trained subjects did not judge the major mode tonic as most suitable. Instead, ratings for the major mode tonic, mediant, dominant, and subdominant did not differ significantly from each other, although together they were rated more highly than were the other tones of the diatonic scale. It would appear, then, that simply hearing a diatonic collection of pitches, independent of their ordering, is not sufficient for reliable key attribution.

Krumhansl and colleagues have elaborated on the role of the diatonic collection in key attribution (see Krumhansl, 1990a, for a review). Subjects listened to a context sequence, which was followed by a probe tone, and they judged how well the probe tone appeared to fit in the context that was provided. The context patterns were configured so as to provide strong cues for key identification; for example, one consisted of a major scale, and another of a sequence of chords that formed a cadence. It was found that the subjects generally gave the tonic the highest rating, and from this it was argued that they were treating the tonic as the most appropriate candidate for a tonal center.

The explanation for the findings of Krumhansl and colleagues has been the subject of considerable debate (see, e.g., Butler, 1989, 1990; Krumhansl, 1990b; West & Fryer, 1990). It was pointed out that in these experiments the tones that were included in the context set were generally given higher ratings than were those that were not so included. Further, there were strong correlations between the number of times a tone appeared in the context set and the strength of its rating by the subject. On such grounds, it was argued that the subjects' ratings could have reflected short-term memory factors rather than knowledge of hierarchies of prominence for tones within keys (see Section V). On the other hand, it could also be argued that the subjects were drawing on their long-term memories for the relative frequencies of occurrence of tones within a key; this could in turn have influenced judgments of hierarchies of prominence. The extent to which ratings were here driven by short-term memory, long-term memory, or yet other factors remains unresolved.

Other theorists have argued that the order in which tones appear is important to key attribution, whether or not the tones are confined to a diatonic collection. As illustration, we can take two musical passages shown in Figure 16. Passage (a) in this figure (G-F#-G-D-C#-D-B-A#-B-G-F#-G) clearly invokes the key of G major, even though two of these tones (C# and A#) are not included in the G-major scale. Passage (b) consists of a haphazard reordering of the tones in passage (a) and clearly does not evoke a well-defined key. Comparison of these two passages indicates that when tones are so ordered as to enable a simple structural description, a key appropriate to this description is readily invoked. However, when the tones are not so ordered, key attribution is more ambiguous.

Another illustration is based on the passage in Figure 19a, which is derived from the example used by Deutsch and Feroe (1981) and is discussed in Deutsch (1984). This passage may be described as an arpeggiation that ascends through the C-major triad (E-G-C) such that each tone of the triad is preceded by a neighbor embellishment, so creating a two-tone pattern. In this way, the key of C major is clearly attributed, despite the fact that two of the tones (D# and F#) are not part of the C-major scale. An algorithm that assigns overwhelming importance to the collection of pitches in the passage would attribute the key of E minor instead. However, if the same sequence of tones is played backward, as in Figure 19b, the key of E minor is now invoked. Other orderings also invoke E minor, even when the passage ends on the note C, as in the example in Figure 19c.

Why, then, is the key of C major so clearly attributed in the pattern shown in Figure 19a, but not in the other patterns? Deutsch (1984) suggested that when presented with a succession of tones, the listener forms groupings based on pitch proximity and, other things being equal, assigns more prominence to the last tone of each grouping. So from the sequence in Figure 19a, the groupings (D#-E), (F#-G), and (B-C) are formed. The tones E, G, and C are then targeted to be combined at a higher level so as to produce an arpeggiation of the C-major triad. In consequence, the key of C major is invoked. However, given the retrograde of the pattern shown in Figure 19b, the groupings (C-B), (G-F#), and (E-D#) are formed



FIGURE 19 Pitch series that illustrate the effects on key attribution of different temporal orderings of a given collection of tones. See text for details. Pattern (a) is derived from Figure 1 of Deutsch and Feroe (1981). (Adapted from Deutsch, 1984.)

instead, and in consequence the tones B, F \sharp , and D \sharp are targeted for higher level grouping. Because these tones form the dominant triad in E minor, and because all the tones in the passage are in the E-minor scale, this key is attributed. On this line of reasoning, the process of key assignment is a complex one, involving such factors as low-level groupings, knowledge of the pitches forming diatonic collections, and knowledge of hierarchies of prominence for tones within different keys.

Bharucha (1984a, 1984b) in parallel hypothesized a similar process that he termed "melodic anchoring." He suggested that a tone becomes "anchored" to its successor if the two are related stepwise along a chromatic or diatonic scale. He further hypothesized that the second tone contributes more than the first one to the listener's sense of key. Bharucha (1984a) provided evidence for this view in a number of experiments.

The importance of order information to key attribution has been stressed particularly by Brown and Butler (see Brown, 1988; Brown & Butler, 1981; Brown, Butler, & Jones, 1994; Butler, 1983, 1992; Butler & Brown, 1984). The authors argued that in attributing a key, listeners rely more on rare than on common intervals, because these correlate less ambiguously with a particular diatonic set, and so provide more reliable information for key identification (Browne, 1981). They also argued that listeners make more accurate key assignments based on rare intervals when these are presented in a temporal order that implies the goal-oriented harmonies that most frequently occur in tonal music. So, for example, when a musical event contains a tritone that is so configured as to imply a dominant-to-tonic progression, the appropriate key will be readily invoked.

Finally, it is evident that other factors are involved in the process of key attribution, such as harmony, rhythm, and meter; however, because of space limitations these factors are not considered here.

IV. PARADOXES BASED ON PITCH CLASS

As described earlier, the pitch of a tone is held to vary along two dimensions: The monotonic dimension of height defines its position along a continuum from low to high, and the circular dimension of pitch class defines its position within the octave (Babbitt, 1960, 1965; Bachem, 1948; Charbonneau & Risset, 1973; Deutsch, 1969, 1972c, 1973b, 1986a; Deutsch & Boulanger, 1984; Forte, 1973; M. Meyer, 1904, 1914; Révész, 1913, Risset, 1969, 1971; Ruckmick, 1929; Shepard, 1964, 1982; Ueda & Ohgushi, 1987).

In order to accommodate the dimensions of pitch class and height in a single spatial representation, it has been suggested that pitch be depicted as a geometrically regular helix, in which the entire structure maps into itself under transposition (Drobisch, 1855; Shepard, 1964, 1965, 1982). Such a representation is shown in Figure 20, and it can be seen that tones that are separated by octaves are depicted as in close spatial proximity. This geometrical model assumes that the dimensions of pitch class and pitch height are orthogonal, so that the pitch class of a tone would not influence its perceived height.

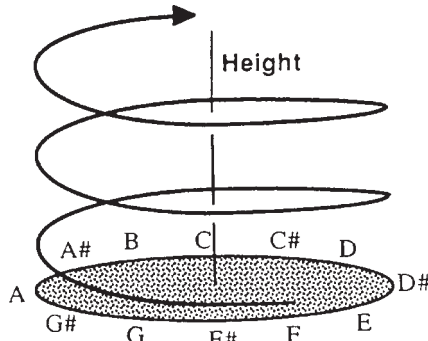


FIGURE 20 Pitch represented as a geometrically regular helix. (Adapted from Shepard, 1965.)

Shepard (1964) noted that the helical model of pitch has an intriguing consequence. If one could suppress the monotonic component of height, leaving only the circular component of pitch class, all tones that were related by octaves could be mapped onto the same tone, which would then have a clearly defined pitch class but an indeterminate height. In this way, the tonal helix would be collapsed into a circle, and judgments of pitch would become completely circular.

In an associated experiment, Shepard generated a set of tones, each of which consisted of 10 sinusoidal components that were separated by octaves and whose amplitudes were determined by a fixed, bell-shaped, spectral envelope. The pitch classes of the tones were varied by shifting the components up or down in log frequency, holding the position and shape of the envelope constant. Shepard argued that because the spectral envelope remained fixed, the perceived heights of these tones would remain constant as their pitch classes were varied.

Subjects listened to ordered pairs of such tones, and they reported in each case whether they heard an ascending or a descending pattern. When the tones within a pair were separated by one or two steps along the pitch class circle (Figure 21) judgments were determined almost entirely by proximity. When the tones were separated by larger distances along the circle, the tendency to follow by proximity gradually lessened, and when they were separated by exactly a half-octave, ascending and descending judgments occurred equally often.

Shepard (1964) concluded from these findings that the dimensions of pitch class and height were indeed orthogonal, arguing that such a view would at all events be expected on common sense grounds:

tonality [i.e., pitch class] seems quite analogous to the attribute of being clockwise or counterclockwise. One of two nearby points on a circle can be said to be clockwise from the other; but it makes no sense to say how clockwise a single point is absolutely.

However, this conclusion does not necessarily follow from Shepard's findings. Where judgments were heavily influenced by proximity, any effect of pitch class on perceived height could have been overwhelmed by this factor. Furthermore, because the data were averaged across pitch classes, any effect of pitch class on

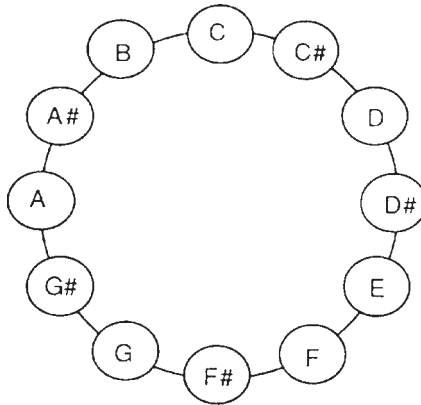


FIGURE 21 The pitch class circle.

perceived height would have been lost in the averaging process. The issue of whether the pitch class of a tone might influence its perceived height therefore remained unresolved in Shepard's study.

A. THE TRITONE PARADOX

Given such considerations, Deutsch (1986a) had subjects listen to ordered pairs of tones that were related by a half-octave (or tritone), and they reported in each case whether they heard an ascending or a descending pattern. Each tone consisted of six octave-related components, whose amplitudes were scaled by a bell-shaped spectral envelope. In order to control for possible effects of the relative amplitudes of the components of the tones, and also to examine the effects of varying their overall heights, the tone pairs were generated under envelopes that were placed at six different positions along the spectrum, which were spaced at half-octave intervals.

It was reasoned that because the tones within each pair were in opposite positions along the pitch class circle, proximity could not here be used as a cue in making judgments of relative height. So if the assumption of orthogonality were correct, these judgments would not be influenced by the pitch classes of the tones. But it was also reasoned that an interaction between the two dimensions might emerge: The listener might perceive tones in one region of the pitch class circle as higher and tones in the opposite region as lower.

More specifically, it was conjectured that listeners might arrange pitch classes as a circular map, similar to a clock face. This map might have a particular orientation with respect to height. For example, C could be in the 12:00 position and F# in the 6:00 position, so that the listener would perceive the tone pairs C-F# (and B-F and C#-D) as descending, and tone pairs F#-C (and F-B and G-C#) as ascending. If, on the other hand, this map were oriented so that F# stood in the 12:00 position

and C in the 6:00 position, the listener would instead perceive the tone pair C-F# as ascending and the pair F#-C as descending.

The hypothesis of an effect of pitch class on perceived height was strikingly confirmed: The judgments of most subjects showed that tones in one region of the pitch class circle were perceived as higher and those in the opposite region as lower. Another striking finding also emerged: The relationship between pitch class and perceived height differed radically from one subject to another. Figure 22 presents, as an example, the judgments of two subjects who showed particularly clear and consistent relationships between pitch class and perceived height. (The judgments were averaged over tones generated under all six spectral envelopes.) The first subject heard tone pairs C#-G, D-G#, D#-A and E-A# as ascending and tone pairs F#-C, G-C#, G#-D, A-D#, A#-E, and B-F as descending. In contrast, the

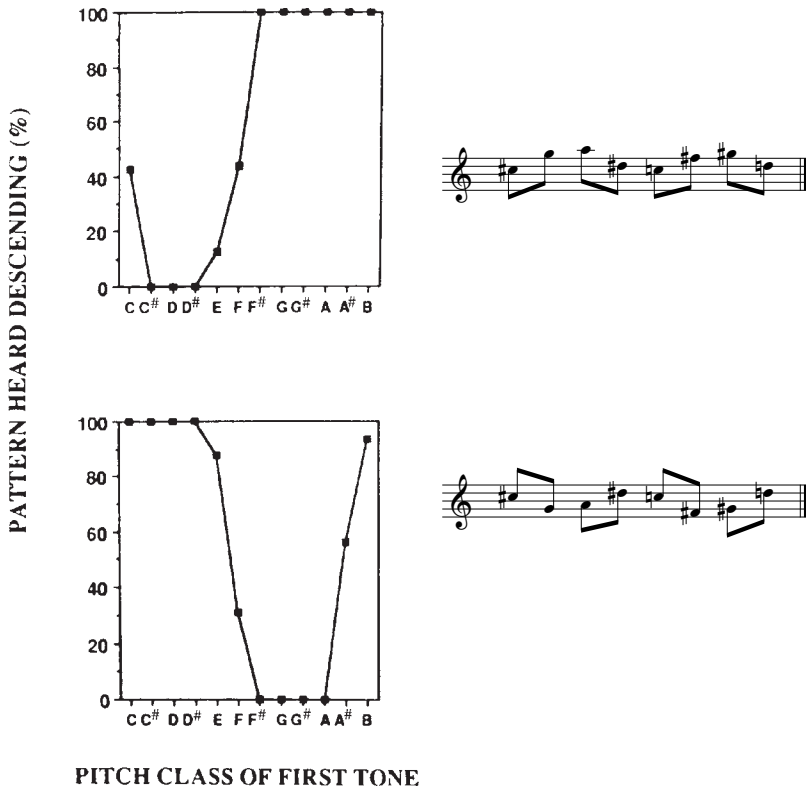


FIGURE 22 The tritone paradox as perceived by two different subjects. The graphs show the percentages of judgments that a tone pair formed a descending pattern, plotted as a function of the pitch class of the first tone of the pair. Notations on the right show how the identical series of tone pairs was perceived by these two subjects. (Data from Deutsch, 1986a.)

second subject heard tone pairs B-F, C-F \sharp , C \sharp -G, D-G \sharp , D \sharp -A, and E-A \sharp as descending and pairs F \sharp -C, G-C \sharp , G \sharp -D, and A-D \sharp as ascending. So for the most part, when the first subject heard an ascending pattern, the second subject heard a descending one, and vice versa. In consequence, also as shown in Figure 22, extended patterns formed of such tone pairs were heard by these two subjects as producing entirely different melodies.

Figure 23 shows the perceptual orientations of the pitch class circle that were derived from the judgments of these two subjects. For the first subject, the peak pitch classes (i.e., those that stood at the highest position along the pitch class circle) were G \sharp and A; however, for the second subject, the peak pitch classes were C \sharp and D instead.

Figure 24 shows the judgments of four more subjects whose patterns were less pronounced than were those shown in Figure 22. These data were taken from experiments in which four spectral envelopes were used, which were spaced at half-octave intervals, and the judgments under these four envelopes were averaged. It can be seen that all four subjects showed clear relationships between pitch class and perceived height; however, the form of the relationship varied from one subject to another. (A set of stimulus patterns comprising a full experiment such as this one appears on the compact disc by Deutsch, 1995).

Deutsch, Kuyper, and Fisher (1987) performed a study to examine perception of the tritone paradox in a general population. A group of subjects were selected on the only criteria that they were undergraduates at the University of California at San Diego, had normal hearing, and could judge reliably whether pairs of sine-wave tones that were related by a tritone formed ascending or descending patterns. The judgments of most of the subjects reflected clear relationships between pitch class and perceived height. Furthermore, computer simulations showed that the patterns obtained in this experiment were extremely unlikely to have occurred by chance. It was concluded that the tritone paradox exists to highly significant extent in this general population.

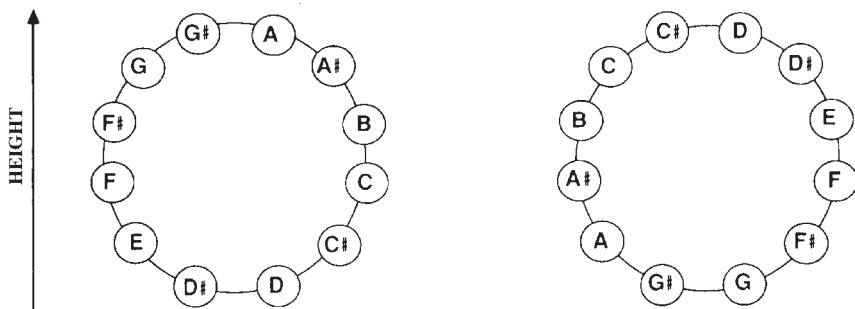


FIGURE 23 Perceptual orientations of the pitch class circle, derived from the judgments of the two subjects whose data are displayed in Figure 22. The circle on the left is derived from the graph shown in the upper portion of Figure 22, and the circle on the right from the graph shown in the lower portion. The pitch classes that mark the highest position along the circle are called peak pitch classes.

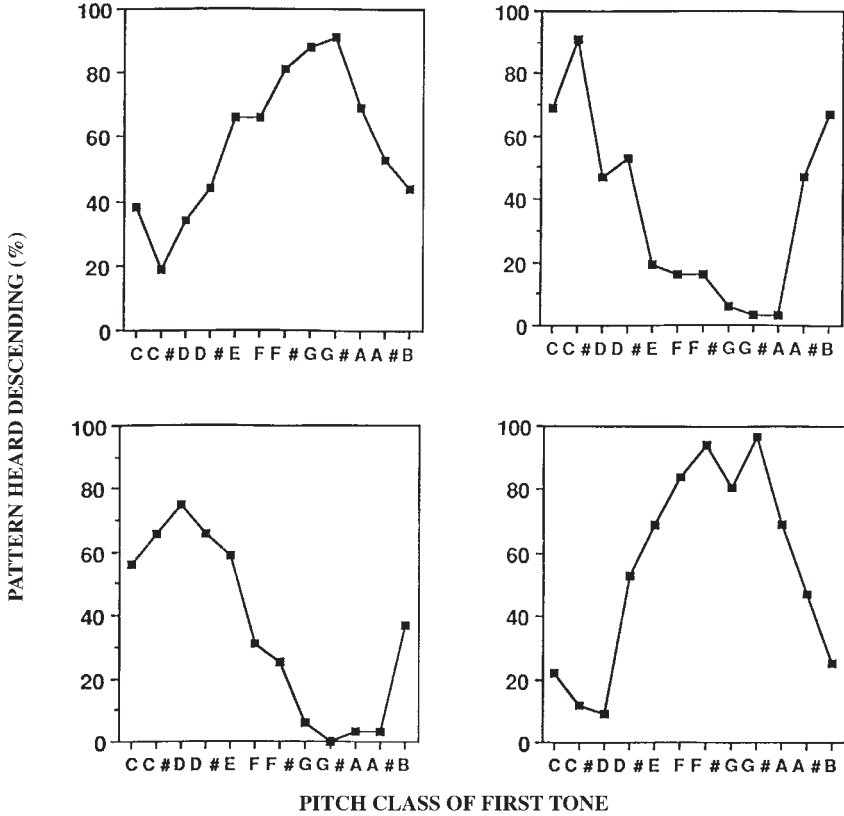


FIGURE 24 The tritone paradox as perceived by four more subjects.

B. THE SEMITONE PARADOX

We next inquire as to what happens when more than one tone is presented at a time. Deutsch (1988b) examined this issue using a basic pattern that consisted of two sequentially presented tone pairs, which were diametrically opposed along the pitch class circle. An example of such a pattern is shown in Figure 25. On one side of the circle the second tone was higher than the first (in this example, G# was followed by A) and on the other side the second tone was lower (in this example, D# was followed by D). In general, subjects linked the tones sequentially in accordance with pitch proximity, so that they perceived the pattern as two stepwise lines that moved in contrary motion. However, the higher line could be heard as ascending and the lower line as descending, or vice versa.

Subjects were presented with such sequential tone pairs, and they judged in each case whether the line that was higher in pitch formed an ascending or a descending pattern. From these judgments it was inferred which pitch classes were heard as higher and which as lower. Taking the tone pairs in Figure 25, for ex-

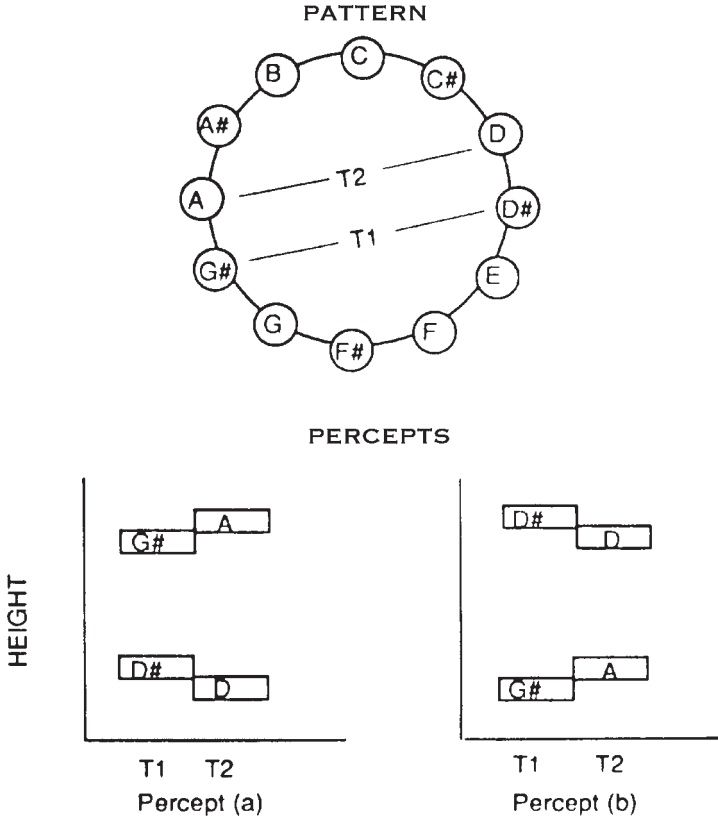
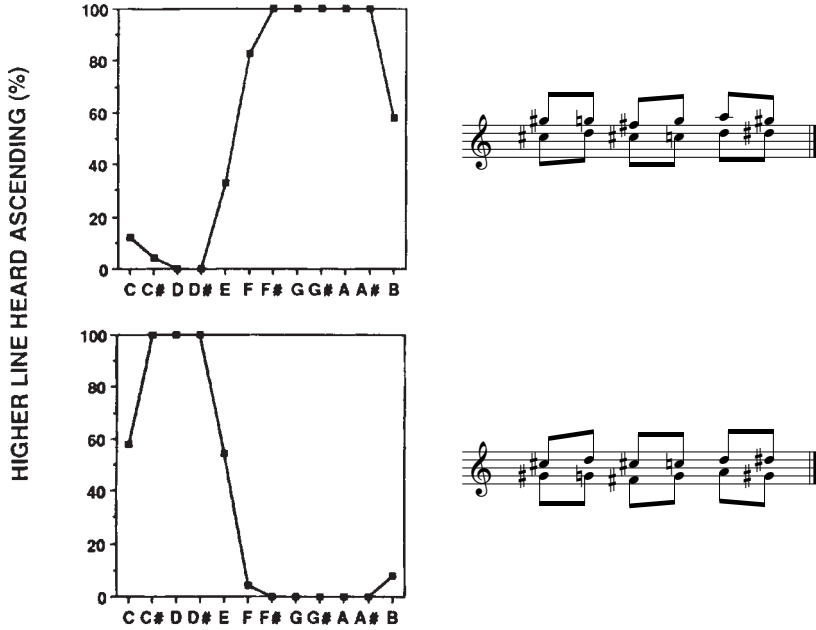


FIGURE 25 Example of pattern giving rise to the semitone paradox, together with two alternative perceptual organizations. Tones G# and D# are presented simultaneously at T1, and tones A and D at T2. Listeners organize this pattern as two stepwise lines that move in contrary motion; that is, they hear the ascending line G#-A together with the descending line D#-D. However, some listeners hear the ascending line as higher [Percept (a)] while other listeners hear the descending line as higher [Percept (b)]. (Adapted with permission from Deutsch, 1988b. ©1988 by The Regents of the University of California.)

ample, if the subject heard the higher line as ascending, this indicated that he or she perceived G# and A as higher and D# and D as lower [as in Percept (a)]. However, if the subject heard the higher line as descending, this indicated that he or she perceived D# and D as higher and G# and A as lower [as in Percept (b)].

Just as with the tritone paradox, subjects' judgments here reflected orderly relationships between the pitch classes of the tones and their perceived heights. Also as with the tritone paradox, the form of this relationship varied radically from one subject to another. This is illustrated in the judgments of two subjects shown in Figure 26. For the first subject, tones F, F#, G, G#, A, and A# were heard as higher and C, C#, D, and D# were heard as lower. In contrast, for the second subject, C#, D,



TONES IN ASCENDING LINE

FIGURE 26 The semitone paradox as perceived by two different subjects. The graphs show 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. For both subjects, when the pattern was transposed, the ascending and descending lines appeared to interchange positions. Moreover, in general, when the first subject heard the higher line as ascending, the second subject heard it as descending, and vice versa. Notations on the right show how the identical series of patterns was perceived by these two subjects. (Reprinted with permission from Deutsch, 1988b. ©1988 by The Regents of the University of California.)

and D# were heard as higher and F, F#, G, G#, A, A#, and B were heard as lower. In consequence, also as shown in Figure 26, musical passages produced by series of such tone pairs were heard by these two subjects in entirely different ways.

C. THE MELODIC PARADOX

We can further ask what happens when more complex patterns are presented. Deutsch, Moore, and Dolson (1986) investigated this question using patterns consisting of three sequentially presented tone pairs. Specifically, the pattern shown in Figure 27 was played in two different keys. In C major, the pattern consisted of the succession of tones D-E-F played together with B-A-G. In F# major, the tones G#-A#-B were played together with E#-D#-C#.

When this pattern was heard unambiguously, the listeners always organized the tones sequentially in accordance with pitch proximity. So they heard one melodic

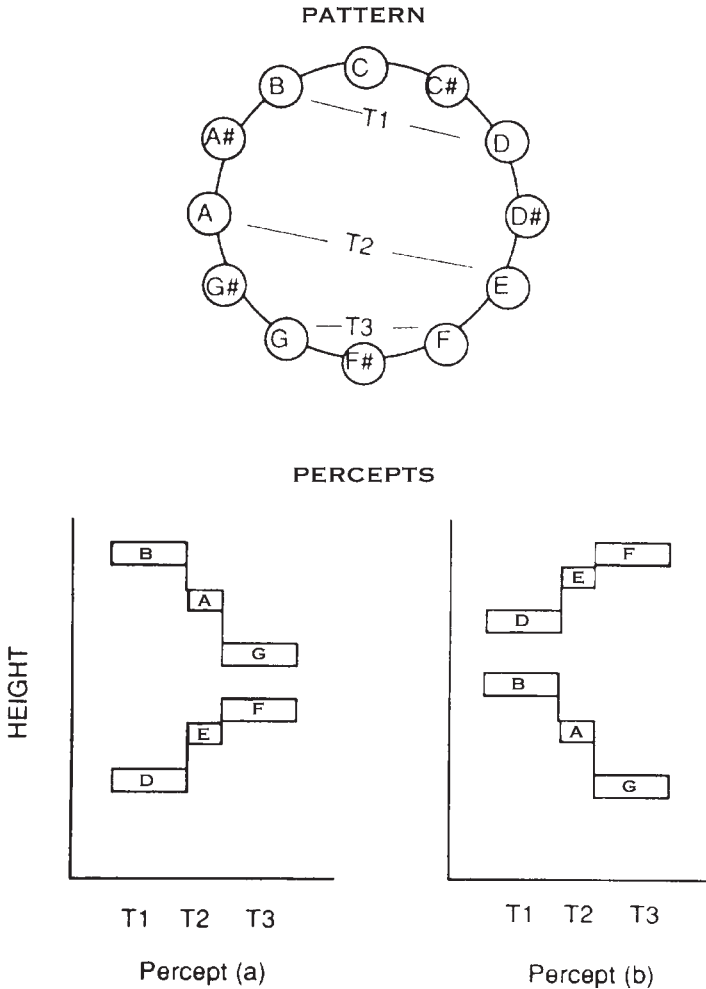


FIGURE 27 Pattern giving rise to the melodic paradox, together with alternative perceptual organizations. Tones D and B are simultaneously presented at time T1, tones E and A are presented at T2, and tones F and G are presented at T3. This pattern is generally heard as two stepwise lines that move in contrary motion. However, some listeners hear the higher line as descending and the lower line as ascending [Percept (a)], whereas others hear the higher line as ascending and the lower line as descending [Percept (b)]. (From experiment by Deutsch et al., 1986.)

line that ascended by a minor third, together with another that descended by a major third. However, also as shown in Figure 27, the descending line could be heard as higher and the ascending line as lower [as in Percept (a)] or the ascending line could be heard as higher and the descending line as lower [as in Percept (b)].

Analogous effects were found to occur here also: When the pattern was played in one key, it was perceived with the higher line ascending. However, when the

pattern was played in the other key, it was heard with the higher line descending instead. So transposing the pattern from one key to the other led to a perceived interchange of voices. Furthermore, when the pattern was played in one of the keys, it was heard with the higher line as ascending by some listeners, but as descending by others.

Deutsch (1988a) performed a further experiment, in which this two-part pattern was played in six different keys: C, D, E, F#, G#, and A# major. The judgments of four subjects were examined, and they all showed orderly effects of key and also differed radically in the direction in which key influenced their judgments. In consequence, extended passages formed of such patterns were heard by these subjects in entirely different ways.

As illustration, Figure 28 displays the percepts of two of the subjects. The first subject heard the pattern in the keys of C and D with the higher line ascending, yet in the keys of E, F#, and G# with the higher line descending. The second subject, in contrast, heard the pattern in the keys of D, E, F#, and G# with the higher line ascending, yet in the keys of C and A# with the higher line descending. Thus for the most part when the first subject heard the higher line ascending, the second subject heard it descending, and vice versa. This is also illustrated in the notation on the right-hand part of the figure.

D. IMPLICATIONS OF THESE MUSICAL PARADOXES

The paradoxes described here show that pitch class and pitch height are not orthogonal dimensions; rather, the perceived height of a tone is systematically related to its position along the pitch class circle, when other factors are controlled for.

The paradoxes are surprising on a number of grounds. First, they provide striking violations of the principle of perceptual equivalence under transposition; a principle that had been assumed to be universal. In the case of the tritone paradox, transposing the pattern from one key to another can cause it to appear to change from an ascending pattern to a descending one, and vice versa. In the case of the paradoxes involving two-part patterns, transposition can result in a perceived interchange of voices.

Another surprising implication concerns absolute pitch, a faculty that is generally assumed to be very rare. Because the majority of listeners experience these musical paradoxes, this means that the majority of us have at least a partial form of absolute pitch, in that we hear tones as higher or lower depending simply on their note names, or pitch classes. It is interesting that other studies have also indicated that absolute pitch is more prevalent than had earlier been assumed, at least in partial form. Terhardt and Ward (1982) and Terhardt and Seewann (1983) found that musicians were able to judge whether or not a passage was played in the correct key, even though they did not possess absolute pitch as conventionally defined. Further, Halpern (1989) reported that when musically untrained subjects were asked to hum the first notes of well-known songs on different occasions, they were remarkably consistent in their choices of pitches from one session to another.

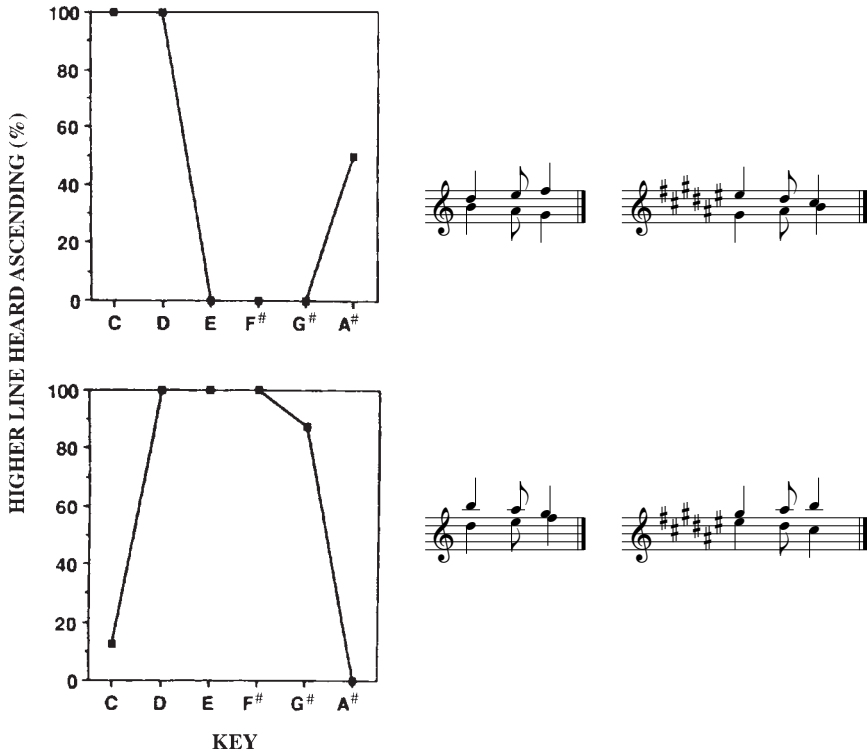


FIGURE 28 Melodic paradox as perceived by two different subjects, when presented in six different keys. In general, as the pattern was transposed, the ascending and descending lines appeared to interchange positions. Moreover, for the most part when the first subject heard the higher line ascending, the second subject heard it descending, and vice versa. Notations on the right show how the identical series of patterns was perceived by these two subjects. (Adapted from Deutsch, 1988a.)

An ingenious study by Levitin (1994) provided further evidence that many people possess a latent form of absolute pitch. He had subjects select a CD that contained a popular song, with which they were very familiar. The subjects were asked to hold the CD, close their eyes, imagine the song, and then reproduce it by singing, humming, or whistling. The first tone that the subject produced was then compared with the equivalent one on the CD. Levitin found that one out of four subjects produced the same pitch (or pitch class) as the equivalent one on the CD, and half the subjects produced a tone whose pitch (or pitch class) was within a semitone of the equivalent. Virtually none of the subjects claimed to have absolute pitch, and the effect appeared to be independent of the subjects' musical background.

Another surprising outcome of this work concerns the striking differences between people in how these paradoxes are perceived. These differences are as pro-

nounced among experienced musicians as among listeners without musical training. It is to the basis of such differences that we now turn.

E. BASIS OF THE TRITONE PARADOX

Studies exploring the bases of these musical paradoxes have focused largely on the tritone paradox. A number of experiments have examined the involvement of spectral factors in this effect. Deutsch (1987) generated such tritone pairs under 12 different spectral envelopes, with peaks spaced at 1/4-octave intervals, so that their positions varied over a three-octave range. Although the relationship between pitch class and perceived height was found sometimes to vary as a function of the overall height of the spectral envelope, and sometimes also as a function of the relative amplitudes of the components of the tones, such variations tended to be small in absolute terms (see also Dawe, Platt, & Welsh, 1998; Deutsch, 1994; Giangrande, 1998).

A number of informal observations led the author to hypothesize that the tritone paradox might be related to the processing of speech sounds. It was conjectured that the listener develops a long-term representation of the overall pitch range of his or her speaking voice. Included in this representation is a specification of the octave band in which the largest proportion of pitch values occurs. The listener then takes the pitch classes delimiting this octave band for speech as marking the highest position along the pitch class circle; this in turn determines his or her orientation of the pitch class circle with respect to height.

In a test of this hypothesis, Deutsch, North, and Ray (1990) selected a group of subjects who showed clear relationships between pitch class and perceived height in making judgments of the tritone paradox. We took a 15-min recording of natural speech from each subject, and from this recording we identified the octave band containing the largest number of pitch values. Comparing across subjects, we obtained a significant correspondence between the pitch classes defining this octave band and those marking the highest position along the pitch class circle, as determined by judgments of the tritone paradox.

Two versions of the hypothesis relating the tritone paradox to the pitch range of speech may then be advanced. The first does not assume that the pitch range of an individual's speech is itself determined by a learned template. The second, and broader, version assumes that we acquire such a template as a result of exposure to speech around us. This template is then used both to constrain our own speech output and also to evaluate the speech of others. If the second hypothesis were correct, we would expect the orientation of the pitch class circle to be similar for people in a given linguistic subculture, but to vary among people in different subcultures.

In a review of the literature concerning the pitch ranges of speech in different linguistic communities, Dolson (1994) described evidence in support of this second hypothesis. First, most people confine the pitch range of their speech to

roughly an octave. Second, within a given linguistic community, in general the speech of females is close to an octave above that of males; for this reason, a template based on pitch class rather than pitch would enable the mapping of male and female speech onto a single mental representation. Further, the pitch ranges of speech differ remarkably little within a given linguistic community (except, of course, for the difference between sexes); however, there are considerable variations in the pitch ranges of speech across different linguistic communities. Moreover, there is a surprising lack of correlation between the pitch range of a person's speech and physiological parameters such as his or her height, weight, chest size, and laryngeal size. This indicates that the pitch range of an individual's speaking voice is based on cultural consensus.

Deutsch (1991) performed a further experiment to test the cultural hypothesis. The judgments of two groups of subjects were compared: The first group had grown up in California and the second group had grown up in the South of England. The two groups were found to differ statistically in their perceptions of the tritone paradox, so that frequently when a Californian subject heard the pattern as ascending, a subject from the South of England heard it as descending, and vice versa (Figure 29).

Other laboratories have obtained further evidence for a geographic association. Giangrande (1998) found that a group of subjects at Florida Atlantic University produced a distribution of peak pitch classes that was similar to the one found by Deutsch (1991) among Californians. Treptoe (1997) found a very similar distribution among subjects at the University of Wisconsin, Steven's Point. In contrast, Dawe et al. (1998) found that a group of students at McMaster University, Ontario, produced a distribution that was quite similar to the one found by Deutsch (1991) for subjects from the South of England.

Examining this correlate in greater detail, Ragozzine and Deutsch (1994) discovered a regional difference in perception of the tritone paradox within the United States. Among subjects who had grown up in the area of Youngstown, Ohio, the perceptions of those whose parents had also grown up in this region differed significantly from those whose parents had grown up elsewhere within the United States. These findings indicate that perception of the tritone paradox is influenced by a template that is acquired in childhood. Further evidence was provided by Deutsch (1996), who found a significant correlation between the way children and their mothers heard the tritone paradox. This correlation was obtained even though the children had all been born and raised in California, whereas their mothers had grown up in many different geographical regions, both within and outside the United States.

We can then ask why such a speech-related template would have evolved. Because the pitch of speech varies with the speaker's emotional state, a template such as hypothesized here could be useful in enabling listeners to make a rapid evaluation of such states. It is also possible that this template could be useful in communicating syntactic aspects of speech.

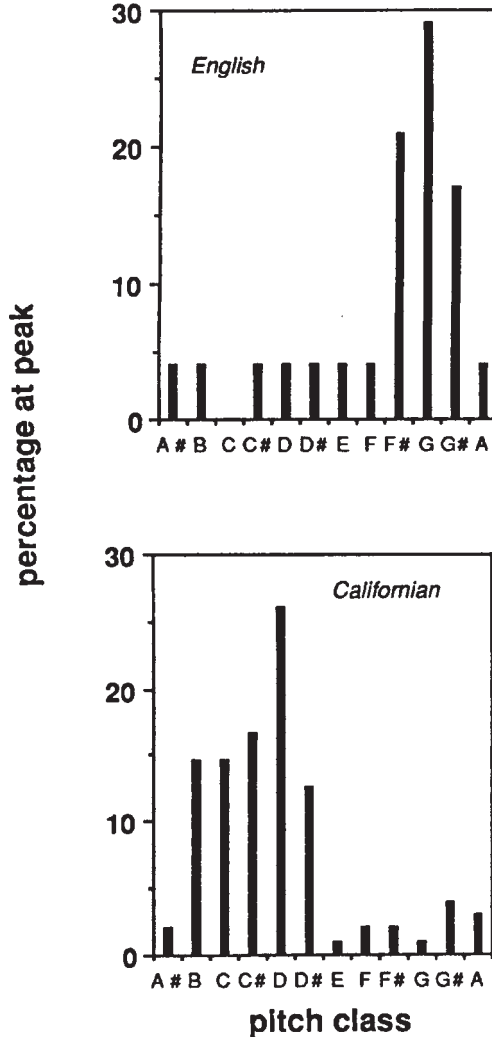


FIGURE 29 Distributions of peak pitch classes in two groups of subjects. The first group had grown up in the south of England and the second group had grown up in California. (Reprinted with permission from Deutsch, 1991. ©1991 by The Regents of the University of California.)

We can then ask whether these perceptual paradoxes occur in natural musical situations. In unpublished experiments, the author has found that the effects persisted with the use of tone complexes whose partials were stretched slightly, so that they did not stand in octave relation. The effects also persisted when the sinusoidal components of the tones were replaced by sawtooth waves, so that the power spectrum of each complex was similar to one produced by several instru-

ments playing simultaneously. Furthermore, imposing a vibrato, a tremolo, or a fast decay such as occurs with a plucked string did not destroy the effects. Most interestingly, the tritone paradox was produced, at least in some individuals, when each tone of a pair consisted of a single harmonic series, with the relative amplitudes of the odd and even harmonics adjusted so that the tones were roughly equivalent in perceived height. It appears, therefore, that effects such as these might well be found in music performed by live instruments, when the composer has introduced ambiguities of height, such as in orchestral pieces by Debussy and Ravel.

V. THE PITCH MEMORY SYSTEM

In this section, we present a theoretical framework for the representation of music in memory. This framework is derived from one originally proposed by Deutsch (1972a, 1975a, 1975d; Deutsch & Feroe 1975), which was based on a series of findings on memory for pitch and for relationships based on pitch. These findings are reviewed, together with more recent findings by the author and other investigators.

It is evident from general considerations that memory for music must be the function of a heterogeneous system, whose various subdivisions differ in the persistence with which they retain information. For example, the system that subserves memory for pitch relationships must be capable of retaining information over very long periods of time, whereas this is not true of the system that retains absolute pitch values. Similarly, the system that retains temporal patterns must preserve information for considerably longer than the system that retains absolute values of duration. Based on such considerations, we can assume that when memory for a musical pattern is tested after different time periods have elapsed, apparent differences in its form of encoding would emerge. (A similar argument concerning memory in general has been made by Craik and Lockhart, 1972.)

More specifically, the model assumes that musical information is initially subjected to a set of perceptual analyses, which are carried out in different subdivisions of the auditory system. Such analyses result in the attribution of values of pitch, loudness, duration, and so on, as well as values of abstracted information, such as harmonic and melodic intervals, durational relationships, and timbres. It is further assumed that in many of these subsystems, information is represented along arrays that are tonotopically organized with respect to a simple dimension, such as pitch or duration, or some higher level dimension such as interval size.

The model further assumes that the outputs of such analyses are projected onto arrays in corresponding subdivisions of the auditory memory system. Thus, for example, one subdivision retains values of pitch, and others retain values of duration, loudness, melodic interval size, and so on. Information is stored in parallel in these different subdivisions; however, the time constants of storage for these sub-

divisions vary considerably. It is further assumed that specific interactions take place within these subdivisions that are analogous to those that occur in systems processing auditory information at the incoming level. The outputs of these different subdivisions then combine during retrieval of information from memory.

Within this framework, a number of different issues may be considered. We can focus on memory for a low-level attribute such as pitch, and inquire concerning how values of this attribute are represented in storage, and also how they are accessed during retrieval. We can also examine how values of different attributes interact with each other in storage and influence each other in retrieval. Here we review the evidence on these issues, focusing largely on the systems underlying memory for pitch and pitch relationships.

A. THE SYSTEM THAT RETAINS ABSOLUTE PITCH VALUES

In considering the characteristics of the system that retains absolute pitch values, a number of hypotheses may be advanced. For example, such memory might simply deteriorate with the passage of time. Another possibility is that pitch information is retained in a general store that is limited in terms of the number of items it can retain, so that memory loss results from a general information overload. As a third possibility, memory for pitch might be the function of an organized system whose elements interact in specific ways.

Let us begin with the following observations. When a tone is presented, and this is followed immediately by another tone that is either identical in pitch to the first or differs by a semitone, most listeners find it very easy to determine whether the two pitches are the same or different. The task continues to be very easy when a silent interval of 6 seconds intervenes between the tones to be compared. Although memory for pitch has been shown to fade gradually with the passage of time (Bachem, 1954; Harris, 1952; Koester, 1945; Rakowski, 1994; Wickelgren, 1966, 1969), the amount of fading during a silent retention interval of 6 sec is so small that it is barely apparent in this situation. However, when eight extra tones are interpolated during the retention interval, the task becomes strikingly difficult; this is true even when the listener is instructed to ignore the interpolated tones. Deutsch (1970a) found that listeners who made no errors in comparing such tone pairs when they were separated by 6 sec of silence made 40% errors when eight tones were interpolated during the retention interval. In a companion experiment, either four, six, or eight tones were interpolated during a retention interval of constant duration, and it was found that the error rate increased with an increase in the number of interpolated tones.

We can conclude that memory for pitch is subject to a small amount of decay with time, and also a large interference effect produced by other tones. What, then, is the basis of this interference effect? One possibility is that the interpolated tones produce attention distraction, and that attention to the tone to be remembered is necessary for memory to be preserved. If this were the case, then the interpolation

of other materials would also result in memory loss, provided that these, too, distracted the listener's attention. As another hypothesis, pitch information might be held in a general store of limited capacity, along with other types of material. Further materials that entered this store would then also impair pitch recognition. As a third hypothesis, pitch information might be retained in a specialized system, and memory loss might result from interactions that occur specifically within this system.

In an experiment to examine these different hypotheses, Deutsch (1970c) had subjects compare the pitches of two tones that were separated by a retention interval of 5 sec duration. The test tones were either identical in pitch or they differed by a semitone. In the first condition, six tones were interpolated during the retention interval (Figure 30). In the second, six spoken numbers were interpolated instead. In both these conditions, listeners were asked to ignore the interpolated materials and simply to judge whether the test tones were the same or different in pitch. A third condition was identical to the second, except that the listeners were asked to recall the numbers in addition to comparing the pitches of the test tones; this ensured that the numbers were attended to and entered memory. In a fourth condition, the subjects were asked simply to recall the numbers.

It was found that the interpolated tones produced a substantial impairment in memory for the pitch of the test tone. However, the interpolated numbers produced only a minimal impairment, even when the subjects were asked to recall them. In addition, the error rate in number recall was no higher when the subjects were simultaneously performing the pitch-recognition task than when they could ignore the test tones. This experiment indicated, therefore, that decrements in pitch memory resulting from interpolated tones are due to interactions that take place within a specialized system.

B. FURTHER EVIDENCE FOR A SEPARATE PITCH MEMORY STORE

Further evidence has been obtained that pitch memory is the function of a specialized system. Deutsch (1974) had subjects compare the pitches of two test tones that were separated by a retention interval that contained eight interpolated tones.

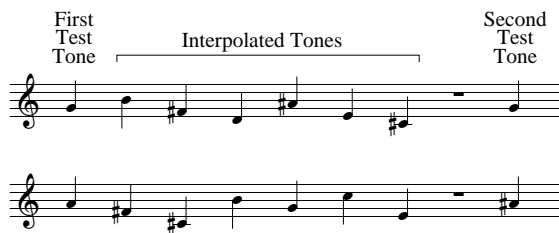


FIGURE 30 Examples of tone series used in experiments to examine the effects of interpolated tones on memory for the pitch of a test tone.

In one condition, the interpolated tones were all drawn from the same octave as the test tones, in a second they were all drawn from the octave above, and in a third they were all drawn from the octave below. The interpolated tones produced substantial interference in all conditions; however, the amount of interference varied depending on the octave in which the interpolated tones were placed. The largest performance decrement occurred when the interpolated tones were in the same octave as the test tones, the next largest when they were in the octave above, and the smallest when they were in the octave below. This experiment indicated, therefore, that the amount of interference produced by interpolated tones depends on the pitch relationships between these and the test tones.

Semal, Demany, and colleagues have shown in a number of experiments that interference in pitch memory results from interactions that take place in a system that is sensitive to pitch relationships, but insensitive to other attributes of sound. In one study, Semal and Demany (1991) had subjects compare the pitches of two test tones that were separated by a retention interval containing six interpolated tones. The test tones were sine waves, and in some conditions, the interpolated tones were also sine waves whereas in others they were of complex spectral composition. Substantial decrements in pitch recognition occurred when the interpolated tones were close in pitch to the test tones, regardless of their spectra. However, when the interpolated tones were remote in pitch from the test tones, the amount of memory impairment they produced was substantially smaller, again regardless of their spectra.

In a further experiment, Semal and Demany (1991) studied the effect of interpolating tones that were composed of several harmonics of a missing fundamental. Again, memory performance depended essentially on the pitches of the interpolated tones and not on their spectral composition. Interpolated tones that were close in pitch to the test tones were associated with poor performance, regardless of their spectra. Performance levels were higher when the interpolated tones were remote in pitch from the test tones, again regardless of their spectra.

In yet another experiment, Semal and Demany (1993) found that differences in the amplitudes of the interpolated tones had remarkably little effect on performance. The amount of memory impairment produced by the interpolated tones was not a monotonically increasing function of their amplitudes, neither did maximal interference occur when the amplitudes of the test and interpolated tones were identical. The authors also found that pitch memory performance was affected very little by whether or not the test and interpolated tones had the same time-varying envelopes.

Semal, Demany, Ueda, and Halle (1995) had subjects make memory judgments concerning words that were spoken at different pitches. The test words were separated by a retention interval during which other materials were interpolated. The interpolated materials were either themselves words that were spoken at different pitches, or they were complex tones that were played at different pitches. The amount of memory impairment produced by the interpolated materials was found to be greater when their pitches were close to those of the test words than when

they were remote from them, regardless of whether words or tones were interpolated. The authors concluded that the pitches of spoken words are not processed in a specialized “speech module,” but rather in a system that is responsible for retaining pitch information, disregarding other attributes of sound.

Other studies have explored the effects on pitch memory of varying the ear of input of the test and interpolated tones. Deutsch (1978c) obtained slightly better memory performance when the test and interpolated tones were presented to different ears rather than to the same ear; however, the difference between these two conditions was fairly small. Kallman, Cameron, Beckstead, and Joyce (1987) confirmed the small advantage produced by delivering the test and interpolated tones to different ears; however, they found that this advantage was present only when the ear of input for the interpolated tones was fixed within a block—the procedure used by Deutsch (1978c). When the ear of input for the interpolated tones varied unpredictably from trial to trial, the advantage disappeared. These experiments provide further evidence that the pitch memory store is remarkably impervious to other attributes of sound.

Finally, Pechmann and Mohr (1992) found that for certain subjects and under certain conditions, information from different memory stores did interfere slightly with pitch recognition. Whereas for musically trained listeners, memory performance was unaffected by the interpolation of spoken words or visual materials, this was not true of musically untrained listeners, who showed small performance decrements when other materials were interpolated. The precise reason for the differences between these two groups of subjects obtained by Pechmann and Mohr remains to be determined.

C. SPECIFIC INTERACTIONS WITHIN THE PITCH MEMORY SYSTEM

We next inquire more specifically into the types of interaction that occur within the system that retains pitch information. In one experiment, Deutsch (1972b) had subjects compare the pitches of two test tones that were separated by a sequence of six interpolated tones. The test tones were either identical in pitch or they differed by a semitone. The effects were explored of placing a tone whose pitch bore a critical relationship to the pitch of the first test tone (the “critical tone”) in the second serial position of the interpolated sequence. This distance varied in steps of $1/6$ tone between identity and a whole-tone separation.

As shown in Figure 31, when the first test tone and the critical tone were identical in pitch, memory facilitation was produced. As the pitch distance between these two tones increased, errors in pitch recognition also increased; they peaked at $2/3$ -tone separation and then decreased, returning to baseline at roughly a whole-tone separation.

Based on these findings, it was conjectured that pitch memory is the function of an array whose elements are activated by tones of specific pitch. These elements are organized tonotopically on a log frequency continuum, and inhibitory interac-

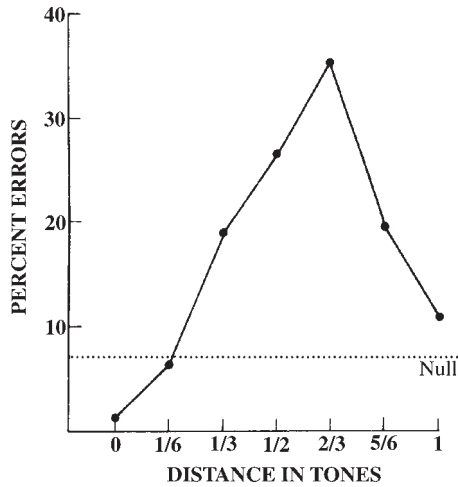


FIGURE 31 Percentages of errors in pitch recognition as a function of the pitch distance between the first test tone and a critical interpolated tone. The line labeled Null shows the error rate in a control condition where all interpolated tones were at least 1-1/2 tones removed in pitch from the first test tone. The maximal error rate occurred when the critical interpolated tone was 2/3 tone removed from the first test tone. (Reprinted with permission from Deutsch, 1972b. ©1972 American Association for the Advancement of Science.)

tions take place along this array that are a function of the distance between the interacting elements. It was further hypothesized that these interactions are analogous to recurrent lateral inhibitory interactions that occur in systems processing sensory information at the incoming level (Ratcliff, 1965). When these elements are inhibited, they produce weaker signals, and this in turn results in increased errors in memory judgment.

A number of observations further support this conjecture. For example, the relative frequency range over which memory disruption occurs corresponds well with the range over which centrally acting lateral inhibition has been found in physiological studies of the auditory system (Klinke, Boerger, & Gruber, 1970). As further evidence, error rates in pitch memory judgment cumulate when two critical tones are interpolated, placed one on either side of the first test tone along the pitch continuum (Deutsch, 1973a). Analogously, lateral inhibitory effects cumulate when two inhibitory stimuli are placed, one on either side of the test stimulus along a visuospatial continuum (Ratcliff, 1965).

If the pitch memory system were indeed organized as a recurrent lateral inhibitory network, one might also expect to find evidence for disinhibition: If a tone that was inhibiting memory for another tone were itself inhibited by a third tone, memory for the first tone should return. Specifically, in sequences where the test tones were identical in pitch, if two critical tones were interpolated, one always 2/3-tone removed from the test tone and the other further removed along the pitch

continuum, errors should vary as a function of the pitch relationship between the two critical tones. The error rate should be highest when these two tones are identical in pitch, decline as the second critical tone moves away from the first, dip maximally at a $2/3$ -tone separation, and then return to baseline. In other words, the curve produced should be roughly the inverse of the curve plotting the original disruptive effect.

To test this prediction, Deutsch and Feroe (1975) performed an experiment in which subjects compared the pitches of two test tones when these were separated by a sequence of six interpolated tones. A tone was always placed in the second serial position of the interpolated sequence whose pitch was $2/3$ tone removed from that of the first test tone; that is, in a relationship expected to produce maximal inhibition. Errors were plotted as a function of the pitch of a second critical tone, which was placed in the fourth serial position, whose pitch relationship to the first critical tone varied in $1/6$ -tone steps between identity and a whole-tone separation.

As can be seen in Figure 32, a systematic return of memory was indeed obtained. The error rate in sequences where the second critical tone was identical in pitch to the first was significantly higher than baseline, and the error rate where the two critical tones were separated by $2/3$ tone was significantly lower than baseline.

A first-order inhibitory function was obtained in a companion experiment, and this was used to calculate the theoretical disinhibition function, assuming that the error rate was determined simply by the strength of the signal produced by the element underlying the first test tone. As also shown in Figure 32, there was a good correspondence between the disinhibition function derived experimentally and the one derived theoretically on the lateral inhibition model. This experiment therefore provided strong evidence that pitch memory elements are indeed arranged as a lateral inhibitory network, analogous to those handling sensory information at the incoming level.

D. ITEM AND ORDER INFORMATION

There is a further effect that causes impairment in pitch memory. When two test tones differ in pitch, and the interpolated sequence includes a critical tone whose pitch is identical to that of the second test tone, increased errors of misrecognition result. This increase is greater when the critical tone is placed early in the interpolated sequence rather than late (Deutsch, 1970b, 1972a; see also Butler & Ward, 1988).

In order to explain this effect, it was conjectured that pitch information is retained both along both a pitch continuum and also along a temporal or order continuum, so as to produce a memory distribution such as shown in Figure 33. According to this conjecture, as time proceeds, this memory distribution spreads along both continua, but particularly along the temporal continuum. As a result of this spread, when the second test tone is presented, the subject recognizes that it

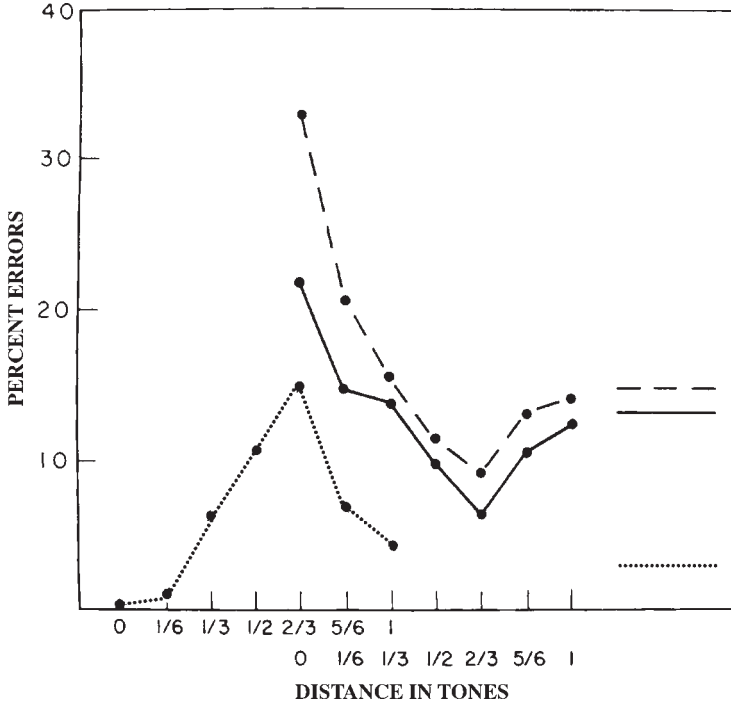


FIGURE 32 Percentages of errors in pitch recognition obtained experimentally and predicted theoretically. Dotted line displays percentage of errors in a baseline experiment that varied the pitch relationship between the first test tone and a critical interpolated tone. (Dotted line at right displays percentage of errors where no tone was interpolated in the critical range.) Solid line displays percentage of errors in an experiment where a tone that was $2/3$ tone removed from the first test tone was always interpolated. Errors are plotted as a function of the pitch relationship between this tone and a second critical interpolated tone that was further removed along the pitch continuum. Dashed line displays percentage of errors for the same experimental conditions predicted theoretically from the lateral inhibition model. (Solid and dashed lines at right display percentages of errors obtained experimentally and predicted theoretically where no further critical tone was interpolated.) (Adapted from Deutsch & Feroe, 1975).

had occurred in the sequence, but is uncertain when it had occurred, and so sometimes erroneously assumes that it had been the first test tone. This hypothesis also would explain the dependence of the effect on the serial position of the critical tone.

A later experiment lent further support to this conjecture. In this experiment, the pitch difference between the first and second test tones was varied, and errors were again plotted as a function of the pitch relationship between the first test tone and the critical tone. It was found, as predicted, that in sequences where the critical tone and the second test tone were on the same side of the first test tone along the pitch continuum, then as the pitch of the second test tone shifted along this contin-

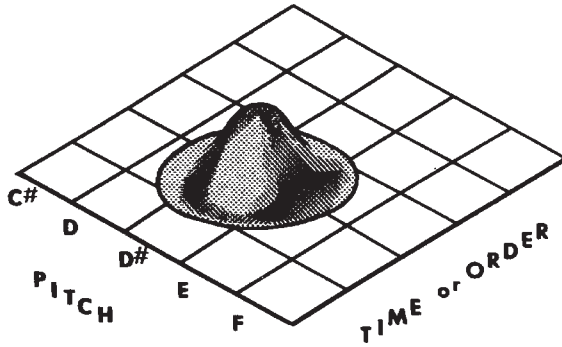


FIGURE 33 Distribution hypothesized to underlie memory for the pitch of a tone. See text for details. (Adapted from Deutsch, 1972a.)

uum, the peak of errors produced by the critical tone shifted in parallel (Deutsch, 1975d).

E. MEMORY ENHANCEMENT THROUGH REPETITION

The conjecture that pitch memory is the function of a temporal or order continuum as well as a pitch continuum gives rise to a further prediction: When the critical tone is identical in pitch to the first test tone, memory performance should be enhanced. More specifically, when the critical tone and the first test tone are identical in pitch, their memory distributions should overlap along the temporal continuum, and this should in turn produce a stronger memory trace for the first test tone (Deutsch, 1972a). This enhancement in performance should be greater when the critical tone is placed early in the interpolated sequence rather than late. Such a pattern of results was indeed obtained in a number of experiments (e.g., Deutsch, 1970b, 1972a, 1975b).

F. OCTAVE GENERALIZATION

As described earlier, there is considerable evidence that pitch is represented along arrays of both pitch height and pitch class. This leads us to inquire whether interference in pitch memory takes place along a pitch height array, or a pitch class array, or whether both such arrays are involved.

In one experiment, Deutsch (1973b) had subjects compare the pitches of two tones that were separated by six interpolated tones. The experiment explored the effects of interpolating tones that bore the same relationship to the test tones as had been found earlier to produce memory disruption, but that were further displaced by an octave. In sequences where the test tones were identical, the effects were studied of including two critical tones, one a semitone higher than the first test tone and the other a semitone lower, but that were further displaced by an

octave. In sequences where the test tones differed, the effects were studied of including a critical tone that was identical in pitch to the second test tone, but again displaced by an octave.

Substantial generalization of disruptive effect resulted from tones that were displaced an octave higher, and a weaker effect from tones that were displaced an octave lower. However, the increase in errors was largest from tones that were placed in the middle octave. It was concluded that disruptive effects in pitch memory take place both along an array of pitch height and also along an array of pitch class.

G. PITCH PROXIMITY AND PITCH MEMORY

Memory for pitch is also influenced by more general relationships between the test and interpolated tones, and among the interpolated tones themselves. For example, in listening to sequences such as we have been describing, the listener processes not only the individual tones, but also the melodic intervals between them. These intervals then provide a framework of pitch relationships to which the test tones can be anchored. So interpolated sequences that form melodic configurations that are more easily processed should be associated with enhanced performance on these tasks.

As described in Chapter 9, there is considerable evidence that melodic patterns are processed more effectively when these are composed of small rather than large intervals, in accordance with the principle of proximity. One might then hypothesize that in our present situation also, interpolated sequences that are composed of small melodic intervals would be associated with higher performance levels than those composed of larger ones. In an experiment to test this hypothesis, Deutsch (1978a) had subjects compare the pitches of two test tones when these were separated by a sequence of six interpolated tones. There were four conditions in the experiment. In the first, the interpolated tones were chosen at random from within a one-octave range, and they were also ordered at random. The second condition was identical to the first, except that the interpolated tones were arranged in monotonically ascending or descending order, so that the average size of the melodic intervals was reduced. In the third condition, the interpolated tones were chosen at random from within a two-octave range, and they were also ordered at random. The fourth condition was identical to the third, except that the interpolated tones were arranged in monotonically ascending or descending order.

As shown in Figure 34, the error rate was found to increase with an increase in the average size of the melodic intervals formed by the interpolated tones. There was no evidence that monotonic ordering of the interpolated tones had any effect, beyond that of producing a smaller average interval size.

It has been shown that there is a striking cross-cultural tendency for the frequency of occurrence of a melodic interval to be inversely correlated with its size (Deutsch, 1978d; Dowling, 1967; Fucks, 1962; Jeffries, 1974; Merriam, 1964; Ortmann, 1926). One might hypothesize that this tendency is based on an increas-

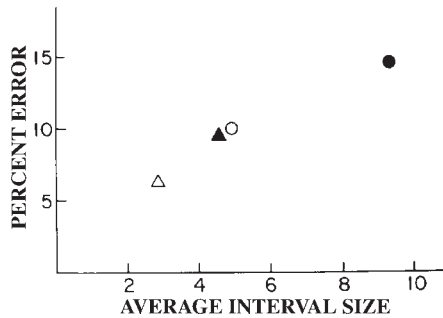


FIGURE 34 Percentages of errors in pitch recognition as a function of the average size of the melodic interval in the sequence. Open triangle: Interpolated tones span a one-octave range and are ordered monotonically. Filled triangle: Interpolated tones span a one-octave range and are ordered at random. Open circle: Interpolated tones span a two-octave range and are ordered monotonically. Filled circle: Interpolated tones span a two-octave range and are ordered at random. (From Deutsch, 1978a.)

ing difficulty in processing melodic intervals as interval size increases. As indicated in the present experiment, this should in turn result in decreased accuracy of pitch recognition judgment.

In a related study, Deutsch (1974) compared the effects on pitch recognition of placing the interpolated tones in different octaves. In the condition where the interpolated tones were drawn from the same octave as the test tones, the error rate was higher than in conditions where they were drawn consistently from the octave above or the octave below. However, the error rate was highest when the interpolated tones were drawn from both the higher and the lower octaves. In this last condition, the interpolated tones formed very large intervals, and as a result, listeners were unable to make use of frameworks of pitch relationships to which the test tones could be anchored. Olson and Hanson (1977) also found that increased distances between the test and interpolated tones were associated with increased errors in pitch recognition (see also Deutsch & Boulanger, 1984).

H. MEMORY FOR TIMBRE

We have presented evidence that the system underlying memory for pitch is subject to specific interactive effects that depend on the pitch relationships between the test and interpolated tones, and that it is remarkably insensitive to other attributes of sound. Starr and Pitt (1997) obtained evidence that the system underlying memory for timbre has analogous characteristics. Their study made use of sets of tones, each of which consisted of a fundamental together with three adjacent harmonics. Thus one type of tone consisted of the fundamental together with harmonics 2, 3, and 4; another consisted of the fundamental together with harmonics 3, 4, and 5; and so on. Eight values of timbre were generated in accordance with this algorithm, and in this way a "spectral similarity dimension" was created. It was found in a preliminary study that tones were judged to be more differ-

ent in timbre as the number of steps separating them along this dimension increased.

Starr and Pitt employed a paradigm that was similar to the one used in the pitch memory studies described earlier. Subjects made comparison judgments between two test tones when these were separated by a retention interval during which extra tones were interpolated. The test tones were either identical in timbre or they differed by three steps along the spectral similarity dimension. The amount of memory impairment was found to increase monotonically as the similarity in timbre between the first test tone and the interpolated tones increased. When the amount of pitch distance between the test and interpolated tones was also varied, the effect of timbre similarity persisted regardless of this pitch distance, and there was only a negligible effect of pitch distance on memory for timbre.

The authors also studied the effects of including a second interpolated timbre in the intervening sequence, and they found that this caused the interference effect of the first interpolated timbre either to increase or to decrease, depending on the distances in timbre involved. They concluded that these findings reflected an effect of disinhibition that is analogous to that found by Deutsch and Feroe (1975) in memory for pitch.

I. MEMORY FOR PITCH RELATIONSHIPS

There is evidence that the system underlying memory for pitch relationships is organized in ways that are similar to the one underlying memory for absolute pitch values. As described earlier, Deutsch (1969) has suggested that interval information is place-coded in the nervous system. Deutsch (1975a) further proposed that this information is projected onto a memory array, such that memory for intervals is the function of a continuum whose elements are activated by the simultaneous or successive presentation of pairs of tones. Tone pairs standing in the same ratio project onto the same elements, and so onto the same point along the continuum; tone pairs standing in closely similar ratios project onto adjacent points along the continuum, and so on. It was further proposed that interactive effects take place along this memory continuum that are analogous to those occurring in the system that retains absolute pitch values. Such effects include memory enhancement through repetition and similarity-based interference.

In considering this issue, informal studies by the author indicated that musically untrained listeners had difficulty in judging directly whether the intervals formed by two pairs of tones were the same or different in size. However, when the same listeners made recognition judgments concerning absolute pitch values, their judgments were strongly influenced by the relational context in which the test tones were placed: They tended to judge test tones as identical in pitch when they were presented in the context of identical harmonic intervals and to judge them as different when they were placed in the context of different intervals (see also Deutsch & Roll, 1974). Given these findings, an experiment was performed to examine indirectly whether memory for harmonic intervals was subject to specific

interactive effects similar to those found in memory for absolute pitch values (Deutsch, 1978b).

Subjects compared the pitches of two test tones when these were both accompanied by tones of lower pitch. The test tones were either identical in pitch or they differed by a semitone. However, the tone accompanying the first test tone was always identical in pitch to the tone accompanying the second test tone. So when the test tones were identical, the intervals formed by the test-tone combinations were also identical. Similarly, when the test tones differed, the intervals formed by the test-tone combinations also differed.

The test-tone combinations were separated by a sequence of six interpolated tones. The tones in the second and fourth serial positions of the interpolated sequence were also accompanied by tones of lower pitch. It was found that when the intervals formed by the interpolated combinations were identical in size to the interval formed by the first test combination, the error rate was lower than when the intervals formed by the interpolated combinations were chosen at random. Furthermore, when the intervals formed by the interpolated combinations differed in size by a semitone from the interval formed by the first test combination, the error rate was higher than when the sizes of the intervals formed by the interpolated combinations were chosen at random.

This experiment indicated that effects analogous to those in the system retaining absolute pitch information also occur in the system retaining abstracted pitch values; namely, memory enhancement through repetition and similarity-based interference.

J. MEMORY FOR DURATION

An experiment by Deutsch (1986b) on memory for duration showed that this has characteristics that are analogous to those found in memory for pitch. Subjects were presented with a pair of blips that defined a first test duration, followed by another pair of blips that defined a second test duration, and they judged whether the second test duration was identical to the first, or longer, or shorter.

In one set of conditions, a series of additional blips was interpolated during the retention interval; these blips defined durations that were in the same range as the test durations. When the interpolated durations were identical to the first test duration, performance levels were higher than when no blips were interpolated. This effect was analogous to the enhancement of performance in pitch memory that occurred when a tone of identical pitch to the first test tone was interpolated. The error rate was also considerably higher when the interpolated blips defined durations that were slightly removed from the first test duration, and judgments reflected distortions in memory for the first test duration in the direction of the interpolated durations. This effect was analogous to the large misrecognition effect that occurred in memory for pitch when a tone that was identical in pitch to the second test tone was interpolated (Deutsch, 1970b, 1972a).

K. INTERACTIONS BETWEEN MEMORY SYSTEMS IN RETRIEVAL

On the hypothesis that memory for different attributes is a function of a number of parallel subsystems, the question arises as to how the outputs of these subsystems influence each other in retrieval. For example, we might expect that judgments of sameness or difference between test tone pitches would be biased by sameness or difference in some other attribute of the tones.

In one experiment, Deutsch and Roll (1974) investigated the effect on pitch memory judgments of manipulating the harmonic relational context in which the test tones were placed. Subjects compared the pitches of two test tones that were both accompanied by tones of lower pitch. The test-tone combinations were separated by a retention interval during which six extra tones were interpolated. In some conditions, the harmonic intervals formed by the test-tone combinations were identical, and in others they differed; these patterns of relationship were present both when the test tones were identical and also when they differed. It was found that relational context had a strong influence on pitch memory judgments. When the test tones were identical, but were placed in different relational contexts, there was an increased tendency for their pitches to be judged as different. Further, when the test tones differed, but were placed in the identical relational context, there was an increased tendency for their pitches to be judged as identical. Further, when the test tones differed, but the test-tone combinations formed intervals that were inversions of each other, there resulted an increase in errors of misrecognition. It was concluded that this misrecognition effect was based on the perceptual equivalence of the inverted intervals (see Section II).

Deutsch (1982) performed an analogous experiment to study the effect of melodic relational context. Subjects compared the pitches of two tones that were each preceded by tones of lower pitch. The test-tone combinations were again separated by a retention interval during which six extra tones were interpolated. A strong effect of melodic context was demonstrated, which was analogous to the effect found for harmonic relational context.

VI. CONCLUSION

In the foregoing pages, we have considered the rules whereby we form abstractions based on pitch, and also how we retain pitch information at different levels of abstraction. Where appropriate, we have considered underlying neurophysiological mechanisms, and we have also drawn on insights provided by music theorists. We have provided evidence that memory for pitch and for low-level pitch relationships is based on a number of highly specialized systems, and that at higher levels pitch information is retained in the form of hierarchies. These findings have contributed to the groundwork on which an understanding of the processing of pitch combinations can be based.

REFERENCES

- Babbitt, M. (1960). Twelve-tone invariants as compositional determinants. *The Musical Quarterly*, 46, 246–259.
- Babbitt, M. (1965). The structure and function of musical theory. *College Music Symposium* 5, 10–21.
- Bachem, A. (1948). Note on Neu's review of the literature on absolute pitch. *Psychological Bulletin*, 45, 161–162.
- Bachem, A. (1954). Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, 26, 751–753.
- Baird, J. W. (1917). Memory for absolute pitch: Studies in psychology. In *Titchener commemorative volume* (p. 69). Worcester, MA.
- Balzano, G. (1980). The group-theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal*, 4, 66–84.
- Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception & Performance*, 6, 501–515.
- Bharucha, J. J. (1984a). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology*, 16, 485–518.
- Bharucha, J. J. (1984b). Event hierarchies, tonal hierarchies, and assimilation: A reply to Deutsch and Dowling. *Journal of Experimental Psychology: General*, 113, 421–425.
- Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist approach. *Music Perception*, 5, 1–30.
- Bharucha, J. J. (1991). Pitch, harmony and neural nets: A psychological perspective. In P. Todd & D. G. Loy (Eds.), *Music and connectionism*. Cambridge, MA: MIT Press.
- Bharucha, J. J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition*, 13, 63–102.
- Blackwell, H. R., & Schlosberg, H. (1943). Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology*, 33, 407–419.
- Boomsblitter, P., & Creel, W. (1961). The long pattern hypothesis in harmony and hearing. *Journal of Music Theory*, 5(2), 2–30.
- Bower, G. H., & Springston, F. (1970). Pauses as recoding points in letter series. *Journal of Experimental Psychology*, 83, 421–430.
- Bower, G., & Winzenz, D. (1969). Group structure, coding and memory for digit series. *Journal of Experimental Psychology*, 80(2, Pt 2), 1–17.
- Bregman, A. S. (1978). The formation of auditory streams. In J. Requin (Ed.), *Attention and performance* (Vol. VII). Hillsdale, NJ: Erlbaum.
- Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception*, 5, 219–250.
- Brown, H., & Butler, D. (1981). Diatonic trichords as minimal cue-cells. *In Theory Only*, 5, 39–55.
- Brown, H., Butler, D., & Jones, M. R. (1994). Musical and temporal influences on key discovery. *Music Perception*, 11, 371–407.
- Browne, R. (1981). Tonal implications of the diatonic set. *In Theory Only*, 5, 3–21.
- Butler, D. (1983). The initial identification of tonal centers in music. In J. Sloboda & D. Rogers (Eds.), *Acquisition of symbolic skills*. New York: Plenum Press.
- Butler, D. (1989). Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception*, 6, 219–241.
- Butler, D. (1990). Reply to Carol Krumhansl. *Music Perception*, 7, 325–338.
- Butler, D. (1992). *The musician's guide to perception and cognition*. New York: Schirmer.
- Butler, D., & Brown, H. (1984). Tonal structure versus function: Studies of the recognition of harmonic motion. *Music Perception*, 2, 6–24.
- Butler, D., & Ward, W. D. (1988). Effacing the memory of musical pitch. *Music Perception*, 5, 251–260.
- Charbonneau, G., & Risset, J. C. (1973). Circularité de jugements de hauteur sonore. *Comptes Rendus de l'Académie des Sciences, Serie B*, 277, 623.

- Collins, A. M., & Quillian, M. R. (1972). How to make a language user. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
- Croonen, W. L. (1994). Effects of length, tonal structure, and contour in the recognition of tone series. *Perception & Psychophysics*, *55*, 623–632.
- Cuddy, L. L., & Cohen, A. J. (1976). Recognition of transposed melodic sequences. *Quarterly Journal of Experimental Psychology*, *28*, 255–270.
- Cuddy, L. L., Cohen, A. L., & Mewhort, D. J. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 869–883.
- Dawe, L. A. Platt, J. R., & Welsh, E. (1998). Spectral-motion aftereffects and the tritone paradox among Canadian subjects. *Perception & Psychophysics*, *60*, 209–220.
- Deese, V., & Grindley, G. C. (1947). The transposition of visual patterns. *British Journal of Psychology*, *37*, 152–163.
- Demany, L., & Armand, P. (1984). The perceptual reality of tone chroma in early infancy. *Journal of the Acoustical Society of America*, *76*, 57–66.
- Deutsch, D. (1969). Music recognition. *Psychological Review*, *76*, 300–307.
- Deutsch, D. (1970a). *The deterioration of pitch information in memory*. Unpublished doctoral dissertation, University of California at San Diego.
- Deutsch, D. (1970b). Dislocation of tones in a musical sequence: A memory illusion. *Nature*, *226*, 286.
- Deutsch, D. (1970c). Tones and numbers: Specificity of interference in short-term memory. *Science*, *168*, 1604–1605.
- Deutsch, D. (1972a). Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology*, *93*, 156–162.
- Deutsch, D. (1972b). Mapping of interactions in the pitch memory store. *Science*, *175*, 1020–1022.
- Deutsch, D. (1972c). Octave generalization and tune recognition. *Perception & Psychophysics*, *11*, 411–412.
- Deutsch, D. (1973a). Interference in memory between tones adjacent in the musical scale. *Journal of Experimental Psychology*, *100*, 228–231.
- Deutsch, D. (1973b). Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics*, *13*, 271–275.
- Deutsch, D. (1974). Generality of interference by tonal stimuli in recognition memory for pitch. *Quarterly Journal of Experimental Psychology*, *26*, 229–234.
- Deutsch, D. (1975a). Auditory memory. *Canadian Journal of Psychology*, *29*, 87–105.
- Deutsch, D. (1975b). Facilitation by repetition in recognition memory for tonal pitch. *Memory & Cognition*, *3*, 263–266.
- Deutsch, D. (1975c). Musical illusions. *Scientific American*, *233*, 92–104.
- Deutsch, D. (1975d). The organization of short-term memory for a single acoustic attribute. In D. Deutsch & J. A. Deutsch (Eds.), *Short-term memory* (pp. 107–151). New York: Academic Press.
- Deutsch, D. (1977). Memory and attention in music. In M. Critchley & R. A. Henson (Eds.), *Music and the brain* (pp. 95–130). London: Heinemann.
- Deutsch, D. (1978a). Delayed pitch comparisons and the principle of proximity. *Perception & Psychophysics*, *23*, 227–230.
- Deutsch, D. (1978b). Interactive effects in memory for harmonic intervals. *Perception & Psychophysics*, *24*, 7–10.
- Deutsch, D. (1978c). Interference in pitch memory as a function of ear of input. *Quarterly Journal of Experimental Psychology*, *30*, 283–287.
- Deutsch, D. (1978d). The psychology of music. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. X, pp. 191–218). New York: Academic Press.
- Deutsch, D. (1979). Octave generalization and the consolidation of melodic information. *Canadian Journal of Psychology*, *33*, 201–205.
- Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, *28*, 381–389.

- Deutsch, D. (1982). The influence of melodic context on pitch recognition judgment. *Perception & Psychophysics*, *31*, 407–410.
- Deutsch, D. (1984). Two issues concerning tonal hierarchies: Comment on Castellano, Bharucha, and Krumhansl. *Journal of Experimental Psychology: General*, *113*, 413–416.
- Deutsch, D. (1986a). A musical paradox. *Music Perception*, *3*, 275–280.
- Deutsch, D. (1986b). Recognition of durations embedded in temporal patterns. *Perception & Psychophysics*, *39*, 179–187.
- Deutsch, D. (1987). The tritone paradox: Effects of spectral variables. *Perception & Psychophysics*, *42*, 563–575.
- Deutsch, D. (1988a). Pitch class and perceived height: Some paradoxes and their implications. In E. Narmour & R. Solie (Eds.), *Explorations in music, the arts, and ideas: Essays in honor of Leonard B. Meyer*. Stuyvesant: Pendragon Press.
- Deutsch, D. (1988b). The semitone paradox. *Music Perception*, *6*, 115–134.
- Deutsch, D. (1991). The tritone paradox: An influence of language on music perception. *Music Perception*, *8*, 335–347.
- Deutsch, D. (1992). Paradoxes of musical pitch. *Scientific American*, *267*, 88–95.
- Deutsch, D. (1994). The tritone paradox: Some further geographical correlates. *Music Perception*, *12*, 125–136.
- Deutsch, D. (1995). *Musical illusions and paradoxes* [CD]. Philomel Records, P. O. Box 12189, La Jolla, CA 92039-2189.
- Deutsch, D. (1996). Mothers and their children hear a musical illusion in strikingly similar ways. *Journal of the Acoustical Society of America*, *99*, 2482.
- Deutsch, D., & Boulanger, R. C. (1984). Octave equivalence and the processing of tonal sequences. *Music Perception*, *3*, 40–51.
- Deutsch, D., & Feroe, J. (1975). Disinhibition in pitch memory. *Perception & Psychophysics*, *17*, 320–324.
- Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, *88*, 503–522.
- Deutsch, D. Kuyper, W. L., & Fisher, Y. (1987). The tritone paradox: its presence and form of distribution in a general population. *Music Perception*, *5*, 79–92.
- Deutsch, D. Moore, F. R., & Dolson, M. (1986). The perceived height of octave-related complexes. *Journal of the Acoustical Society of America*, *80*, 1346–1353.
- Deutsch, D., North, T., & Ray, L. (1990). The tritone paradox: Correlate with the listener's vocal range for speech. *Music Perception*, *7*, 371–384.
- Deutsch, D., & Roll, P. L. (1974). Error patterns in delayed pitch comparison as a function of relational context. *Journal of Experimental Psychology*, *103*, 1027–1034.
- Dewar, K. M., Cuddy, C. L., & Mewhort, D. J. K. (1977). Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 60–67.
- Dewitt, L. A., & Crowder, R. G. (1986). Recognition of novel melodies after brief delays. *Music Perception*, *3*, 259–274.
- Dibben, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *Music Perception*, *12*, 1–25.
- Divenyi, P. L., & Hirsh, I. J. (1978). Some figural properties of auditory patterns. *Journal of the Acoustical Society of America*, *64*, 1369–1385.
- Dolson, M. (1994). The pitch of speech as function of linguistic community. *Music Perception*, *11*, 321–331.
- Dowling, W. J. (1967). *Rhythmic fission and the perceptual organization of tone sequences*. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.
- Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics*, *4*, 37–40.
- Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 342–354.

- Dowling, W. J. (1986). Context effects on melody recognition: Scale-step and interval representation. *Music Perception*, 3, 281–296.
- Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, 50, 305–313.
- Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 49, 524–531.
- Drobisch, M. (1855). Über musikalische Tonbestimmung und Temperatur. In *Abhandlungen der Königlich sächsischen Gesellschaft der Wissenschaften zu Leipzig. Vierter Band: Abhandlungen der mathematisch-physischen Classe. Zweiter Band* (pp. 3–121). Leipzig: S. Hirzel.
- Edelman, G. M., Gall, W. E., & Cowan, W. M. (1988). *Auditory function: Neurobiological bases of hearing*. New York: Wiley.
- Edworthy, J. (1985). Melodic contour and musical structure. In P. Howell, I. Cross, & R. J. West (Eds.), *Musical structure and cognition* (pp. 169–188). Orlando, FL: Academic Press.
- Ernst, G. W., & Newell, A. (1969). *GPS: A case study in generality and problem solving*. New York: Academic Press.
- Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, DC: Winston.
- Forte, A. (1973). *The structure of atonal music*. New Haven, Connecticut: Yale University Press.
- Francès, R. (1988). *The perception of music* (W. J. Dowling, Trans.). Hillsdale, NJ: Erlbaum. (Original work published 1958)
- Fucks, W. (1962). Mathematical analysis of the formal structure of music. *Institute of Radio Engineers Transactions on Information Theory*, 8, 225–228.
- Gagné, R. M. (1962). The acquisition of knowledge. *Psychological Review*, 67, 355–365.
- Giangrande, J. (1998). The tritone paradox: Effects of pitch class and position of the spectral envelope. *Music Perception*, 13, 253–264.
- Greeno, J. G., & Simon, H. A. (1974). Processes for sequence production. *Psychological Review*, 81, 187–196.
- Hall, D. (1974). Quantitative evaluation of musical scale tunings. *American Journal of Physics*, 48, 543–552.
- Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, 17, 572–581.
- Handel, S. (1973). Temporal segmentation of repeating auditory patterns. *Journal of Experimental Psychology*, 101, 46–54.
- Hanson, A. R., & Riseman, E. M. (Eds.) (1978). *Computer vision systems*. New York: Academic Press.
- Harris, J. D. (1952). The decline of pitch discrimination with time. *Journal of Experimental Psychology*, 43, 96–99.
- Helmholtz, H. von. (1954). *On the sensations of tone as a physiological basis for the theory of music* (2nd English ed.) New York: Dover. (Original work published 1859)
- Helmholtz, H. von. (1844). The origin of the correct interpretations of our sensory impressions. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 7, 81–96.
- Hershman, D. P. (1994). Key distance effects in ecological contexts. In I. Deliège (Ed.), *Proceedings of the 3rd International Conference on Music Perception and Cognition, Liège* (pp. 243–244). Liège: ICMPC.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106–154.
- Humphreys, L. F. (1939). Generalization as a function of method of reinforcement. *Journal of Experimental Psychology*, 25, 361–372.
- Idson, W. L., & Massaro, D. W. (1978). A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics*, 24, 551–565.
- Janata, P., & Reisberg, D. (1988). Response-time measures as a means of exploring tonal hierarchies. *Music Perception*, 6, 161–172.
- Jeffries, T. B. (1974). Relationship of interval frequency count to ratings of melodic intervals. *Journal of Experimental Psychology*, 102, 903–905.

- Jones, M. R. (1974). Cognitive representations of serial patterns. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, NJ: Erlbaum.
- Jones, M. R. (1978). Auditory patterns: Studies in the perception of structure. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Vol. VIII. Perceptual coding*. New York: Academic Press.
- Julesz, B., & Hirsh, I. J. (1972). Visual and auditory perception: An essay of comparison. In E. E. David & P. B. Denes (Eds.), *Human communication: A unified view*. New York: McGraw-Hill.
- Kallman, H. J., Cameron, P. A., Beckstead, J. W., & Joyce, E. (1987). Ear of input as a determinant of pitch-memory interference. *Memory & Cognition*, *15*, 454–460.
- Kallman, H. J., & Massaro, D. W. (1979). Tone chroma is functional in melody recognition. *Perception & Psychophysics*, *26*, 32–36.
- Klinke, R., Boerger, G., & Gruber, J. (1970). The influence of the frequency relation in dichotic stimulation upon the cochlear nucleus activity. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Sijthoff: Leiden.
- Knudsen, E. I., du Lac, S., & Esterly, S. D. (1987). Computational maps in the brain. *Annual Review of Neuroscience*, *10*, 41–65.
- Koester, T. (1945). The time error in pitch and loudness discrimination as a function of time interval and stimulus level [Special issue]. *Archives of Psychology*, 297.
- Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt.
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, *11*, 346–374.
- Krumhansl, C. L. (1990a). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Krumhansl, C. L. (1990b). Tonal hierarchies and rare intervals in music cognition. *Music Perception*, *7*, 309–324.
- Krumhansl, C. L., Bharucha, J. J., & Kessler, E. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 24–36.
- Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory & Cognition*, *11*, 325–334.
- Krumhansl, C. L., Sandell, G. J., & Sergeant, D. C. (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception*, *5*, 31–78.
- Leewenberg, E. L. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, *84*, 307–349.
- Lerdahl, F. (1988). Tonal pitch space. *Music Perception*, *5*, 315–349.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Levitin, D. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, *56*, 414–423.
- Longuet-Higgins, H. C. (1962a). Letter to a musical friend. *Music Review*, *23*, 244–248.
- Longuet-Higgins, H. C. (1962b). Second letter to a musical friend. *Music Review*, *23*, 271–280.
- Longuet-Higgins, H. C. (1978). The perception of music. *Interdisciplinary Science Reviews*, *3*, 148–156.
- Mach, E. (1959). *The analysis of sensations and the relation of the physical to the psychical* (C. M. Williams, Trans.; W. Waterlow, review and supplement). New York: Dover. (Original work published in German, 1906)
- McLean, R. S., & Gregg, L. W. (1967). Effects of induced chunking on temporal aspects of serial retention. *Journal of Experimental Psychology*, *74*, 455–459.
- Merriam, A. P. (1964). *The anthropology of music*. Evanston, IL: Northwestern University Press.
- Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.
- Meyer, L. B. (1973). *Explaining music: Essays and explorations*. Berkeley, CA: University of California Press.
- Meyer, M. (1904). On the attributes of the sensations. *Psychological Review*, *11*, 83–103.
- Meyer, M. (1914). Review of G. Révész, "Zur Grundlegency der Tonpsychologie." *Psychological Bulletin*, *11*, 349–352.

- Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. *Handbook of Mathematical Psychology*, 2, 419–493.
- Miller, G. A., Galanter, E. H., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.
- Mueller, G. E., & Schumann, F. (1894). Experimentelle Beiträge zur Untersuchung des Gedächtnisses. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 6, 81–190, 257–339.
- Narmour, E. (1990). *The analysis and cognition of basic melodic structures*. Chicago: University of Chicago Press.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353–383.
- Nettl, B. (1956). *Music in primitive culture*. Cambridge, MA: Harvard University Press.
- Olson, R. K., & Hanson, V. (1977). Interference effects in tone memory. *Memory & Cognition*, 5, 32–40.
- Ortmann, O. (1926). On the melodic relativity of tones [Special issue]. *Psychological Monographs*, 35, Whole No. 162.
- Oura, Y. (1991). Constructing a representation of a melody: Transforming melodic segments into reduced pitch patterns operated on by modifiers. *Music Perception*, 9, 251–266.
- Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, 9, 441–474.
- Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory & Cognition*, 20, 314–320.
- Pickler, A. G. (1966). Logarithmic frequency systems. *Journal of the Acoustical Society of America*, 39, 1102–1110.
- Piston, W. (1948). *Harmony* (2nd ed.) London: Norton.
- Pitts, W., & McCulloch, W. S. (1947). How we know universals: The perception of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 9, 127–147.
- Plomp, R., Wagenaar, W. A., & Mimpfen, A. M. (1973). Musical interval recognition with simultaneous tones. *Acustica*, 29, 101–109.
- Ragozzine, F., & Deutsch, D. (1994). A regional difference in perception of the tritone paradox within the United States. *Music Perception*, 12, 213–225.
- Rakowski, A. (1994). Investigating short term auditory memory with the method of vocal pitch control. In *Proceedings of the Stockholm Music Acoustics Conference* (pp. 53–57).
- Ratliff, F. (1965). *Mach bands: Quantitative studies of neural networks in the retina*. San Francisco, CA: Holden Day.
- Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review*, 77, 481–495.
- Restle, F. (1972). Serial patterns: The role of phrasing. *Journal of Experimental Psychology*, 92, 385–390.
- Restle, F., & Brown, E. (1970). Organization of serial pattern learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 4, pp. 249–331). New York: Academic Press.
- Révész, G. (1913). *Zur grundlegency der tonpsychologie*. Leipzig: Feit.
- Risset, J.-C. (1969). Pitch control and pitch paradoxes demonstrated with computer-synthesized sounds. *Journal of the Acoustical Society of America*, 46, 88(A).
- Risset, J.-C. (1971). Paradoxes de hauteur: Le concept de hauteur sonore n'est pas le meme pour tout le monde. *Proceedings of the Seventh International Congress on Acoustics, Budapest, 1970*, 613–616.
- Rosner, B. S., & Meyer, L. B. (1982). Melodic processes and the perception of music. In D. Deutsch (Ed.), *The psychology of music* (1st ed.). New York: Academic Press.
- Rosner, B. S., & Meyer, L. B. (1986). The perceptual roles of melodic process, contour, and form. *Music Perception*, 4, 1–40.
- Ruckmick, C. A. (1929). A new classification of tonal qualities. *Psychological Review*, 36, 172–180.
- Scandura, J. M. (1970). Role of rules in behavior: Toward an operational definition of what (rule) is learned. *Psychological Review*, 77, 516–533.

- Schenker, H. (1956). *Neue musikalische theorien und phantasien: Der Freie Satz*. Vienna: Universal Edition.
- Schmuckler, M. A. (1989). Expectation in music: investigations of melodic and harmonic processes. *Music Perception*, 7, 109–150.
- Schoenberg, A. (1951). *Style and idea*. London: Williams & Norgate.
- Schoenberg, A. (1967). In G. Strong (Ed.), *Fundamentals of musical composition*. New York: St. Martin's Press.
- Schoenberg, A. (1969). *Structural functions of harmony* (rev. ed.) New York: Norton. (Original work published 1954)
- Schreiner, C. E. (1992). Functional organization of the auditory cortex: Maps and mechanisms. *Current Opinion in Neurobiology*, 2, 516–521.
- Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *Journal of the Acoustical Society of America*, 89, 2404–2410
- Semal, C., & Demany, L. (1993). Further evidence for an autonomous processing of pitch in auditory short-term memory. *Journal of the Acoustical Society of America*, 94, 1315–1322.
- Semal, C., Demany, L., Ueda, K., & Halle, P. (1995). Short-term pitch memory: Speech versus non-speech. In: *Proceedings of the meeting of the International Society for Psychophysics, Cassis*. Cassis: International Society for Psychophysics.
- Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 36, 2345–2353.
- Shepard, R. N. (1965). Approximation to uniform gradients of generalization by monotone transformations of scale. In D. L. Mostofsky (Ed.), *Stimulus generalization*. Stanford, CA: Stanford University Press.
- Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (1st ed., pp. 343–390). New York: Academic Press.
- Simon, H. A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 79, 369–382.
- Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review*, 70, 534–546.
- Simon, H. A., & Sumner, R. K. (1968). Pattern in music. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: Wiley.
- Starr, G. E., & Pitt, M. A. (1997). Interference effects in short-term memory for timbre. *Journal of the Acoustical Society of America*, 102, 486–494.
- Suga, N. (1990). Cortical computational maps for auditory imaging. *Neural Networks*, 3, 3–21.
- Sutter, M. I., & Schreiner, C. E. (1991). Physiology and topography of neurons with multi-peaked tuning curves in cat primary auditory cortex. *Journal of Neurophysiology*, 65, 1207–1226.
- Takeuchi, A. H. (1994). More on key-distance effects in melody recognition: A response to van Egmond and Povel. *Music Perception*, 12, 143–146.
- Takeuchi, A. H., & Hulse, S. H. (1992). Key-distance effects in melody recognition reexamined. *Music Perception*, 10, 1–24.
- Terhardt, E., & Seewann, M. (1983). Aural key identification and its relationship to absolute pitch. *Music Perception*, 1, 63–83.
- Terhardt, E., & Ward, W. D. (1982). Recognition of musical key: Exploratory study. *Journal of the Acoustical Society of America*, 72, 26–33.
- Thomson, W. (1991). *Schoenberg's error*. Philadelphia: University of Pennsylvania Press.
- Trainor, L. L., & Trehub, S. E. (1993). Musical context effects in infants and adults: Key distance. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 615–626.
- Treptoe, R. (1997). *Further study of the tritone paradox*. Unpublished manuscript, University of Wisconsin–Stevens Point.
- Ueda, K., & Ohgushi, K. (1987). Perceptual components of pitch: Spatial representation using multidimensional scaling technique. *Journal of the Acoustical Society of America*, 82, 1193–1200.
- Van Egmond, R., & Povel, D. J. (1994a). Factors in the recognition of transposed melodies: A comment on Takeuchi and Hulse. *Music Perception*, 12, 137–142.

- Van Egmond, R., & Povel, D. J. (1994b). Similarity judgments of transposed melodies as a function of overlap and key distance. In I. Deliège (Ed.), *Proceedings of the 3rd International Conference on Music Perception and Cognition, Liège* (pp. 219–220). Liège: ICMP.
- Van Egmond, R., & Povel, D. J. (1996). Perceived similarity of exact and inexact transpositions. *Acta Psychologica*, 92, 283–295.
- Van Egmond, R., Povel, D. J., & Maris, E. (in press). The influence of height and key on the perceptual similarity of transposed melodies. *Perception & Psychophysics*.
- Van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Unpublished doctoral thesis, Technische Hogeschool, Eindhoven, Holland.
- Vicario, G. B. (1983). *Micro- and macromelodies*. Paper presented at the Fourth Workshop on Physical and Neuropsychological Foundations of Music, Ossiach.
- Vitz, P. C., & Todd, T. C. (1967). A model of learning for simple repeating binary patterns. *Journal of Experimental Psychology*, 75, 108–117.
- Vitz, P. C., & Todd, T. C. (1969). A coded element model of the perceptual processing of sequential stimuli. *Psychological Review*, 76, 433–449.
- Von Ehrenfels, C. (1890). Über Gestaltqualitäten. *Vierteljahrsschrift für Wissenschaftliche Philosophie*, 14, 249–292.
- Warren, R. M., & Warren, R. P. (1968). *Helmholtz on perception: Its physiology and development*. New York: Wiley.
- Weber, G. (1824). *Versuch einer Heordeneten Theorie*. Mainz: B. Schotts Sohne.
- Werner, H. (1925). Über Mikromelodik und Mikroharmonik. *Zeitschrift für Psychologie*, 98, 74–89.
- Werts, D. (1983). *A theory of scale references*. Unpublished doctoral dissertation, Princeton University, Princeton, NJ.
- West, R. J., & Fryer, R. (1990). Ratings of suitability of probe tones as tonics after random orderings of notes of the diatonic scale. *Music Perception*, 7, 253–258.
- Wever, E. G., & Bray, C. W. (1930). The nature of the acoustic response: The relation between sound frequency and frequency of impulses in the auditory nerve. *Journal of Experimental Psychology*, 13, 373–387.
- White, B. (1960). Recognition of distorted melodies. *American Journal of Psychology*, 73, 100–107.
- Wickelgren, W. A. (1966). Consolidation and retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, 72, 250–259.
- Wickelgren, W. A. (1967). Rehearsal grouping and the hierarchical organization of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology*, 19, 97–102.
- Wickelgren, W. A. (1969). Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology*, 6, 13–61.
- Winston, P. H. (1973). Learning to identify toy block structures. In R. L. Solo (Ed.), *Contemporary issues in cognitive psychology: The Loyola Symposium*. Washington, DC: Winston.
- Yngve, V. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 104, 444–466.

