9

The Processing of Pitch Combinations

Diana Deutsch

I. INTRODUCTION

In this chapter the processing of pitch combinations is examined at several levels. First, we inquire into the types of abstraction that give rise to the perception of local features. Such features may be considered analogous to those of orientation or angle size in vision. We have 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

ISBN 0-12-213560-1

as contour. Next, we consider how combinations of such features are abstracted so as to give rise to perceptual equivalences and similarities. We then examine processing where these higher level abstractions are themselves combined according to various rules.

Investigations into mechanisms of visual shape perception have led to a distinction between an early process, where many low-level abstractions are passively carried out in parallel, and a later process, where questions are asked of these low-level abstractions based on hypotheses about the scene to be analyzed (Hanson & Riseman, 1978). This distinction between abstractions that are formed passively from "bottom-up" and those that result from a "top-down" process is important in music also. As we shall see, much of musical shape analysis occurs only when the context is such as to allow for the confirmation of expectations.

The final sections of the chapter are concerned with memory. It is clear that a musical sequence is retained simultaneously at different levels of abstraction and that the information at these different levels combines to influence memory judgments. Interactions occurring within the different memory systems are examined, as are the ways the outputs of these systems interact during retrieval.

II. FEATURE ABSTRACTION

A. Octave Equivalence

It is clear that a strong perceptual similarity exists between tones separated by octaves—that is, whose fundamental frequencies stand in the ratio of 2:1 (or a power of 2:1). 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 is placed (G_3, F_4, etc.). Octave duplications also occur in the scales of other cultures (Nettl, 1956).

Various experimental observations related to octave equivalence have been reported. Baird (1917) and Bachem (1954) have both found that listeners with absolute pitch may sometimes place a note in the wrong octave, even though they name it correctly. Experiments using conditioning procedures have demonstrated generalization of response to tones separated by octaves, both in people (Humphreys, 1939) and in animals (Blackwell & Schlosberg, 1943). Interference effects in memory for pitch have also been shown to exhibit octave generalization (Deutsch, 1973a).

Given the perceptual similarity between tones separated by octaves, it has been suggested that pitch be treated as a bidimensional attribute: the first dimension representing overall pitch level or *tone height*, and the second defining the position of a tone within the octave, or *tone chroma* (Meyer, 1904, 1914; Révész, 1913; Ruckmick, 1929; Bachem, 1948; Shepard, 1964). Contemporary music theorists make an analogous distinction between *pitch* and *pitch class* (Babbitt, 1961, 1965; Forte, 1973).

B. Interval and Chord Equivalence

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 size when the fundamental frequencies of their components stand in the same ratio. This principle forms an important basis for 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 separated by the same number of semitones are given the same name, such as major third, minor sixth, and so on. Contemporary music theory also treats tone pairs separated by the same number of semitones as perceptually equivalent.

Chords consisting of three or more simultaneous tones are also classified in part on the basis of the frequency ratios of their components. However, a simple listing of these ratios is not sufficient to define a chord. For example, a major triad and a minor triad both have as their components a major third (five semitones) and a minor third (four semitones) and a fifth (seven semitones). Thus, the fact that the minor third lies above the major third in the one instance and below it in the other is of perceptual importance, and needs to be taken into account in considering how chord abstraction might be achieved by the nervous system.

Given the principles of octave equivalence and interval equivalence, one might hypothesize that intervals whose components are placed in different octaves are also perceptually equivalent. This assumption is frequently made by contemporary music theorists who refer to such intervals as in the same *interval class*. However, traditional music theory assumes this equivalence for simultaneous but not for successive intervals. Simultaneous intervals whose components have reversed position by being placed in different octaves are termed inversions (Piston, 1948). Thus, a simultaneous interval of *n* semitones is considered perceptually equivalent to a simultaneous interval of 12-n semitones.

Laboratory evidence for the perceptual similarity of inverted intervals has been obtained. Plomp, Wagenaar, and Mimpen (1973) required subjects to identify intervals formed by simultaneous tone pairs, and they found that confusions occurred between intervals that were inversions of each other. Further evidence was provided by Deutsch and Roll (1974) and is discussed below.

For the case of intervals formed by successive tone pairs, the experimental evidence is complicated. As will be discussed below, it appears that interval class perception here occurs only indirectly through a process of hypothesis confirmation, where the features directly apprehended are pitch class and interval.

C. Proposed Physiological Substrates

Following Drobisch (1846, 1855), various psychologists have suggested that the phenomenon of the tone chroma (or pitch class) be accommodated by representing pitch as a helix, with tones separated by octaves lying most proximal within each turn of the helix. One can take this suggestion literally and propose that such a mapping of pitch exists in the auditory system, so that columns are formed of neural units that respond to tones spaced at octave intervals. Unfortunately, there is no physiological evidence for such columnar organization. On the other hand, units have been found that exhibit peaks of sensitivity at octave intervals. Such units could mediate the perceptual equivalence of tones standing in octave relation, and this hypothesis is described in detail below.

Various models for the perceptual equivalence of intervals and chords have been advanced. Pitts and McCulloch (1947) proposed that the auditory cortex is composed of layers, each layer containing a topographic projection of frequency-specific units. In each projection, units responding to frequencies 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. They further hypothesized the existence of fibers that traverse this columnar mass parallel to each other in a slantwise direction. Three such slanting pathways would therefore define a three-note chord. Such a mechanism could mediate transposition of simultaneous intervals and chords but could not mediate transposition of successive intervals nor the perceptual similarity of intervals and chords related by inversion.

Another mechanism, suggested by Boomsliter and Creel (1961), was based on the volley theory of pitch perception (Wever & Bray, 1930). They pointed out that when the components of two frequency combinations stand in the same ratio, these combinations should generate the same pattern of firing. That is, one pattern would be produced by the ratio 2:3, another by the ratio 4:5, and so on. They therefore proposed that the perceptual equivalence of simultaneous intervals is mediated by recognition of these patterns. This model would require an ability on the part of the nervous system to follow frequencies at much higher rates than has been established (Deutsch & Deutsch, 1973). A further difficulty for the model is that it cannot account for the perceptual equivalence of successive intervals and chords nor for the similarity of inverted intervals and chords.

Deutsch (1969) proposed a mechanism for the abstraction of first-order pitch relationships that accommodates both octave equivalence and also the equivalence of intervals and chords under inversion. This hypothesis was modeled on findings concerning the abstraction of low-order specific features by the visual system, such as orientation and angle size. It appears that such abstractions are accomplished in several stages (Hubel & Wiesel, 1962). Units with circular receptive fields[1] appear to project onto higher order units in such a way that units whose receptive fields taken together form straight lines converge onto the same higher order units. These higher order units respond to lines of specific orientation presented in a specific position in the visual field. It further appears that these higher order units project onto yet higher order units in such a way that those units responding to lines presented in a given

[1] The receptive field of a unit is that region which, when stimulated, causes a change in the activity of this unit.

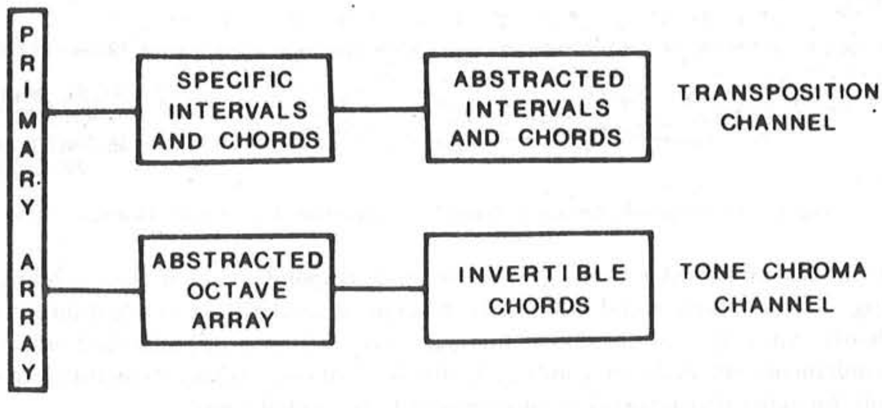

Fig. 1. Flow diagram for abstraction of pitch relationships. It is assumed that pitch information is simultaneously abstracted along two parallel channels, one mediating transposition phenomena and the other mediating octave equivalence effects (from Deutsch, 1969).

orientation but in different positions in the visual field converge onto the same unit. These units therefore respond to lines of specific orientation presented in different positions in the visual field.

The proposed mechanism for abstraction of first-order pitch relationships consists of two parallel channels along each of which information is abstracted in two stages (Fig. 1). Channel A mediates the perceptual equivalence of intervals and chords under transposition. In the first stage of abstraction along this channel, first-order units responding to tones of specific pitch are linked in groups of two and three to second-order units. These units therefore respond to specific intervals and chords (such as the combination of C_4, F_4 and G_4, or of 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 these second-order units are linked to third-order units in such a way that all units that are activated by tones standing in the same relationship are linked together. So, for example, all units activated by an ascending interval of four semitones (a major third) converge onto one unit, all units activated by a descending interval of seven semitones (a perfect fifth) converge onto another unit, all units activated by a major triad onto another unit, and so on (Fig. 2).[2]

Channel B mediates the equivalence of tones separated by octaves (tone chroma or pitch class). In the first stage of abstraction along this channel, first-order units responding to tones of specific pitch are linked in such a way that units underlying tones separated by octaves converge onto the same second-order unit. These units therefore respond to tones in a given pitch class regardless of actual pitch. In the second stage of abstraction, these second-order units converge in groups of two and

[2] Only intervals and chords formed out of elements of the 12-tone chromatic scale are described here, for the purpose of clarity. However, it is assumed that first-order units responding to tones that are not elements of the 12-tone scale are also linked to higher order units in this fashion. This theory therefore makes no assumptions about temperament.

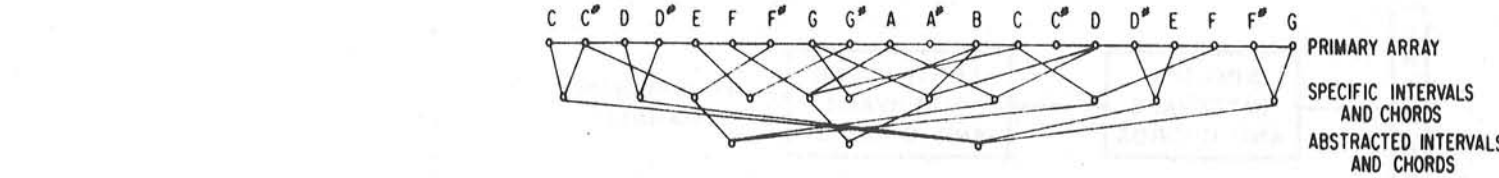

Fig. 2. Two stages of abstraction along the transposition channel (from Deutsch, 1969).

three onto third-order units, which therefore respond to pitch class combinations (Fig. 3). Such units would mediate the perceptual similarity of inverted intervals and chords. Since it is assumed that interval class is directly apprehended only where simultaneous intervals are concerned, 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, some relevant findings may be cited. Suga, O'Neill, and Manabe (1979) describe neurons in the auditory cortex of the bat that showed facilitation when the first harmonic of a tone was delivered simultaneously with the second harmonic so that the combination formed a perfect fifth. Other units showed facilitation when the first and third harmonics were simultaneously presented, so that the combination formed an octave; yet others showed facilitation when the second and third harmonics were simultaneously presented, so that the combination formed a fourth. Such units often responded poorly to single tones in isolation but strongly and consistently when the appropriate tonal combination was presented. On the above model, units with such characteristics are hypothesized to occur at the first stage of abstraction along Channel A, i.e., the channel mediating interval and chord perception.

Evans (1974) reports the existence of neurons in the auditory cortex of the cat that exhibit peaks of sensitivity at more than one band of frequencies. Peaks spaced at octave intervals were commonly found. Also Suga and Jen (1976) note the presence of

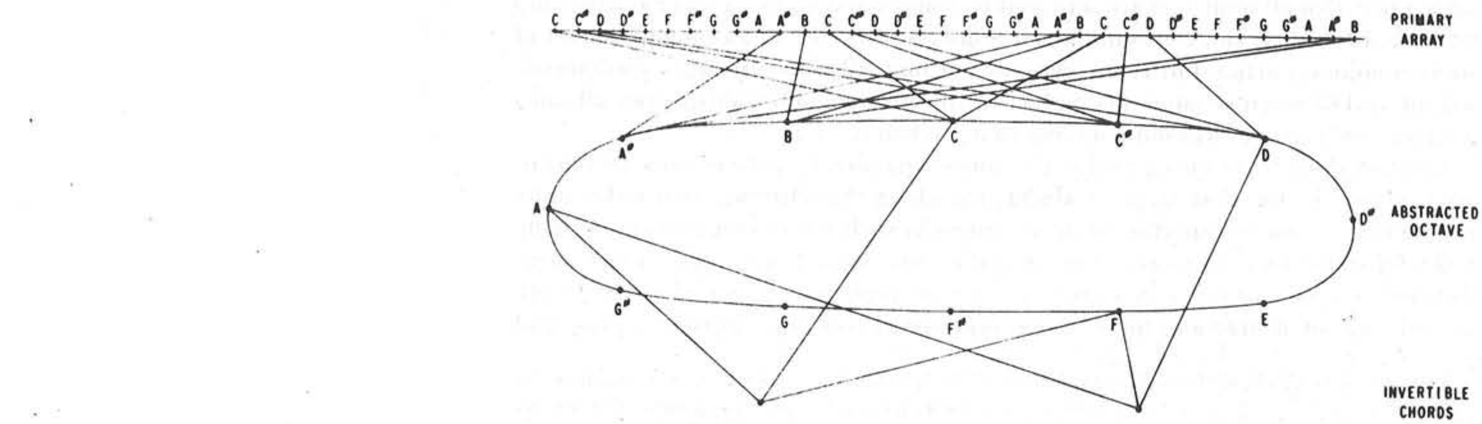

Fig. 3. Two stages of abstraction along the channel mediating octave equivalence effects (from Deutsch, 1969).

neurons in the auditory cortex of the bat that showed two peaks of sensitivity that were approximately harmonically related. They write: "since no neurons with a double peaked tuning curve were found at the periphery . . . it is evident that harmonically related components in acoustic signals are converging on some single neurons at higher levels in the auditory system." In the above model, units with these characteristics were hypothesized to occur at the first level of abstraction on Channel B, i.e., as mediating octave equivalence for single tones.

Units with other patterns of facilitation were also found in these neurophysiological studies. This is hardly surprising since the cat and the bat would not be expected to have the same mechanisms for abstraction of pitch relationships as humans. However, it is interesting that the two classes of neuron hypothesized on the model have been shown to exist, i.e., those with multiple peaks for single tones, and those that require the presentation of specific tones in combination to be securely activated. Furthermore, examples of neurons with characteristics as specifically hypothesized on the model were uncovered in these studies.

D. Contour

In recognizing a segment of music, we employ global as well as specific cues. These include, for example, overall pitch range, the distribution of interval sizes, the proportion of ascending versus descending intervals, and so on. The use of global cues has been best documented for the case of contour in the processing of linear sequence. As shown in the examples in Fig. 4 (taken from Schoenberg, 1967) melodies can be represented by their distinctive contours. We appear to be very sensitive to such information. For example, Werner (1925) found that listeners were able to recognize familiar melodies when these were transformed onto very small scales, so that the intervals were grossly distorted in size. Later, White (1960) found that listeners were able to recognize melodies to some extent when all the intervals were arbitrarily set to one semitone, so that the interval information was entirely removed, apart from the directions of pitch change. When the relative sizes of the intervals were retained, even though their absolute sizes were altered, performance was considerably enhanced. Recent studies by Dowling (1978), Dowling and Fujitani (1971), Idson and Massaro (1978), and Kallman and Massaro (1979) have confirmed and extended such findings.

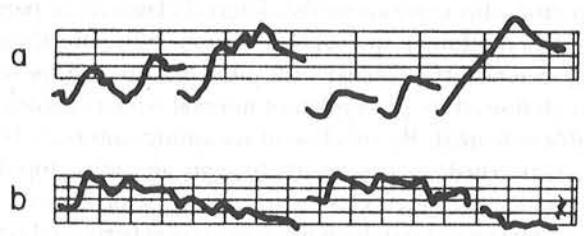

Fig. 4. Contours from Beethoven piano sonatas as represented by Schoenberg. a. from Sonata in C minor, Op. 10/1-III, m 1-8. b. from Sonata in D, Op. 10/3-III, m 1-16 (from Schoenberg, 1967).

E. Interval Class

If 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 and D_5 in combination form the same interval class as G_2 and F_6. Whether interval class identity gives rise to perceptual equivalence is a matter for debate. As mentioned above, experimental evidence for such equivalence has been found where simultaneous intervals are concerned (Plomp *et al.*, 1973; Deutsch & Roll, 1974). Further compelling evidence is provided by the fact that we easily recognize root progressions of chords as abstractions.

Where successive intervals are concerned, however, the issue is complicated. If interval class were indeed a perceptual invariant, we should experience no difficulty in recognizing a melody when its component tones are placed in different octaves. This issue was examined experimentally by Deutsch (1972a). The first half of the tune "Yankee Doodle" was generated under various conditions. First, it was produced without transformation in each of three adjacent octaves. Then, it was generated such that each tone was in its correct position within the octave, but the choice of octave placement varied randomly between these same three octaves. And finally, the tune was generated as a series of clicks so that the pitch information was entirely removed but the rhythmic information retained. This was to provide a measure of identification performance on the basis of rhythm alone.

The different versions of the tune were played to separate groups of subjects, who were given no clues to its identity besides being assured that it was well known. Although the untransformed versions were universally recognized, recognition of the randomized octaves version was no better than for the version where the pitch information was removed entirely. However, when the subjects were later informed of the identity of the tune, and were again presented with the randomized octaves version, they now found they were able to follow the tune to a large extent. They were thus able to use octave generalization to *confirm* the identity of the tune, even though they had been unable to *recognize* it in the absence of prior information. It was concluded that this confirmation was achieved by the listeners' imagining the tune simultaneously with hearing the randomized octaves version. In this way they could match each note as it arrived with their auditory image and so confirm that the two were indeed in the same pitch class.

It would appear from this experiment that interval class can be perceived, but not as a first-order abstraction. Rather, perception occurs indirectly through a process of hypothesis testing in which the listener uses pitch class to transpose each tone to the appropriate octave, followed by perception of interval which enables the hypothesis to be confirmed or disconfirmed. By this line of reasoning, interval class, where successive intervals are concerned, is perceived through an active top-down process, in contrast with interval and pitch-class perception, which result from a passive bottom-up process. The extent to which interval class is perceived would then depend critically on the expectations of the listener.

Deutsch examined this issue again (1976, 1979), using a short-term recognition paradigm. Listeners were presented with a standard six-tone melody, followed by a comparison melody. The comparison was always transposed up four semitones from the standard. On half of the trials, this transposition was exact so that the set of intervals and their orders were preserved. On the other half, two of the tones in the transposed melody were permuted. The permuted tones were always a semitone apart in pitch, and no tones were permuted that were either at the beginning or the end of the melody or that were adjacent to each other. Thus, in the permuted sequences, four out of the five successive intervals were changed in size by a semitone; however, the contour of the melody (as defined by the sequence of directions of pitch change) was unaltered regardless of whether or not the comparison melody was an exact transposition of the standard. (This invariance of contour was necessary to insure that contour could not be used as a basis for recognition judgments.)

There were four conditions in the experiment. In the first condition the standard melody was presented once, followed by the comparison melody. In the second condition the standard melody was repeated six times before presentation of the comparison melody. Here, all repetitions were exact. In the third condition the standard melody was again repeated six times, but now on half of the repetitions the melody was transposed in its entirety an octave higher and on the other half it was transposed an octave lower. In the fourth condition the standard melody was again repeated six times, but now on each repetition the individual tones in the melody were placed alternately in the higher and the lower octaves.

The results of the experiment are shown on Fig. 5. It can be seen that exact repetition of the melody produced a substantial improvement in comparison perfor-

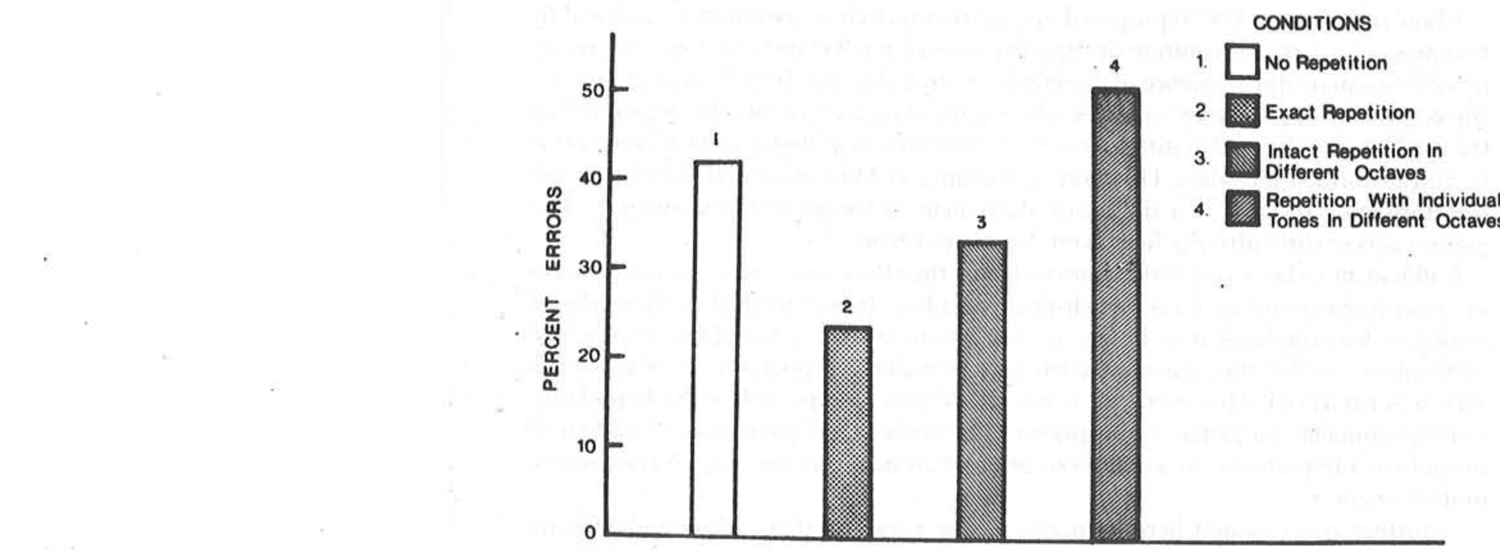

Fig. 5. Percent errors in different conditions of experiment studying the effects of octave displacement on consolidation of melodic information (from Deutsch, 1979).

mance, and an improvement also occurred when the melody was repeated intact in the higher and lower octaves. However, when the melody was repeated such that its component tones alternated between the higher and lower octaves, performance was significantly poorer than when the melody was not repeated at all. This experiment again strongly indicates that interval class cannot be treated as a first-order perceptual feature. Repetition of a set of successive intervals resulted in consolidation of memory for these intervals; however, repetition of a set of interval classes did not do so.

These findings are as expected on the model of Deutsch (1969), since this model assumes that the linkages that give rise to abstraction of successive intervals occur between units underlying specific pitches, and not between higher order units underlying pitch classes. Idson and Massaro (1978) proposed an alternative explanation. They argued that the placement of the tones in different octaves resulted in an alteration of melodic contour, and that this in turn acted to disrupt recognition performance. As evidence for this they point out that when the component tones of melodies were placed in different octaves but contour was preserved, recognition of these melodies was at a higher level than when contour was not preserved (Dowling & Hollombe, 1977; Idson & Massaro, 1978). However, preservation of contour would be expected to improve performance regardless of how interval calss is perceived, since contour alone has been shown to be a powerful cue in melody recognition (Werner, 1925; White, 1960, Dowling, 1978). Once the listener has guessed the identity of the melody on the basis of contour (or any other cue for that matter), he can then confirm or disconfirm his hypothesis by a process of matching each tone as it arrives with its octave equivalent (Deutsch, 1972a, 1978a). The findings on the preservation of contour are not, therefore, evidence against an explanation in terms of Deutsch's (1969) model.

Idson and Massaro (1978) proposed instead that melody recognition is mediated by two processes: first, recognition of the succession of pitch classes, and second, recognition of contour (the sequence of directions of pitch change). If their hypothesis were correct, then recognition of melodies where pitch class and contour are preserved but the component tones are in different octaves should be at as high a level as recognition of untransformed melodies. However, Kallman and Massaro (1979) found that this transformation resulted in a significant decrement in recognition performance. This poses a severe difficulty for Idson and Massaro's theory.

Kallman and Massaro (1979) also investigated the effect on recognition performance of preserving contour under octave displacement but altering pitch class by randomly raising or lowering each tone by one or two semitones. They found that recognition performance under this transformation was considerably poorer than where pitch class was preserved. However, this result would also be expected on the hypothesis-testing argument, for just as a perception of the correct set of pitch classes would tend to confirm a hypothesis, so would perception of an incorrect set of pitch classes tend to disconfirm it.

A further point should here be made. In the experiment by Idson and Massaro (1978), subjects were given the names of a small set of test melodies, and they were presented with these melodies under various transformations for hundreds of trials, allowing ample opportunity for hypothesis-testing. On the other hand Deutsch (1972a) and Kallman and Massaro (1979) presented subjects with each melody only once and did not inform them of the identities of these melodies so that hypothesis testing was much more difficult. In comparing the results from these two types of paradigm, we find that recognition performance where pitch class was preserved but octave placement was randomized was considerably better in the study of Idson and Massaro than in the other two studies. This difference would be expected on the theory that listeners recognize interval class through the mediation of hypothesis-testing, but provides a further difficulty for Idson and Massaro's model.

In this regard it is instructive to consider the use of octave jumps in traditional music. If the present line of reasoning is correct, such jumps can be made with impunity provided the musical setting is such that the displaced tone is expected by the listener. We should therefore suppose octave jumps to be limited to such situations. Indeed, this appears to be the case. In one such situation a melodic line is presented several times without transformation. A clear set of expectations having thus been established, a jump to a different octave occurs. The melodic line on Fig. 6a, for instance, occurs after this line has been presented several times with no octave jumps. Another such situation is where the harmonic structure is clear and unambiguous, so that again the displaced tones are highly probable to the listener. This is illustrated in the segment shown on Fig. 6b. It should also be observed that when octave jumps occur in traditional music, often the identical pitch class is quickly repeated in the new octave. This also occurs in the segment in Fig. 6b. According to the model of Deutsch (1969), this pitch class identity should be recognized directly. The repeated pitch class then provides a means of placing the tones in the new octave in correct relationship with the tones in the previous octave.

The technique of 12-tone composition employs very frequent octave jumps (see below). This raises the question of whether the listener does in fact recognize as perceptually equivalent two presentations of the same tone row under octave displacement. Given the evidence and arguments outlined above, such recognition

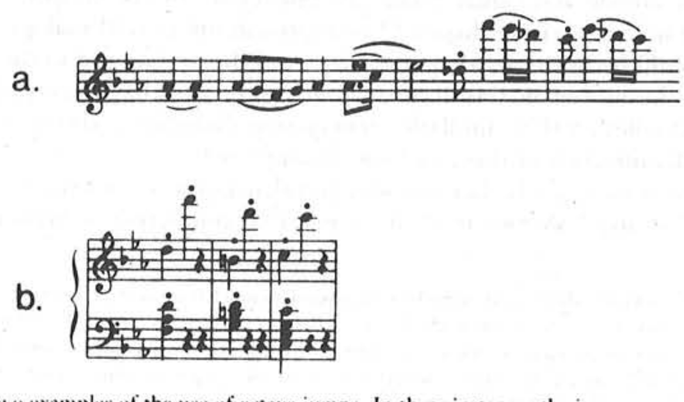

Fig. 6. Two examples of the use of octave jumps. In these instances the jumps are readily processed. a. from Beethoven, *Rondo in C, Op. 5, No 1*. b. from Beethoven, *Sonata in C minor, Op. 10, No. 1*.

should be possible in principle, but only if the listener is very familiar with the material, or if its structure is such as to arouse strong expectations (see also Meyer, 1973).

III. HIGHER ORDER ABSTRACTIONS

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

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

> A tonal series occurs in 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 (Van Noorden, 1975; Deutsch, 1975a; Bregman, 1978; Divenyi & Hirsh, 1978). In the visual representation of the sequence shown on Fig. 11 of Chapter 4, for example, the principle of proximity emerges clearly. We may therefore inquire whether auditory analogues also exist for perceptual equivalences found in vision (Julesz & Hirsh, 1972).

Von Ehrenfels (1890) in his influential paper on form perception, pointed out that a melody when transposed retains its essential form, the *Gestaltqualität,* provided that the relations among individual tones are unaltered. In this respect, he argued, melodies are similar to visual shapes. On our present intermodal analogy, transposing a melody would be like translating a shape to a different location in the visual field. Shapes may be moved in this way without destroying their perceptual identities (Deese & Grindley, 1947). Similarly, transposing melodies to different pitch levels may leave identification of these melodies unimpaired.[3]

We may next inquire whether further equivalences can be demonstrated for musical shapes that are analogous to their visuospatial counterparts. Schoenberg (1951)

[3]This is clear from everyday experience where long-term memory is concerned (Deutsch, 1969; Attneave & Olson, 1971) and also occurs in some short-term situations (Divenyi & Hirsh, 1978). However, short-term recognition of transposed melodies may be difficult (Attneave & Olson, 1971; Cuddy & Cohen, 1976; Dowling, 1978). This probably reflects the projection of interval information onto highly overlearned unequal-interval scales (Deutsch and Feroe, 1981).

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 subconciously 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 such that all ascending intervals become descending intervals and vice versa ("inversion")[4], when it is presented in reverse order ("retrogression"), or when it is transformed by both these operations ("retrograde-inversion"). Figure 7 illustrates Schoenberg's use of his theory in compositional practice. As he 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. That is, he assumed that octave displacement would not destroy the perceptual identity of a musical configuration. His assumptions of perceptual equivalence under transposition, retrogression, inversion, and octave displacement are fundamental to the theory of 12-tone composition (Babbitt, 1961, 1965). This compositional technique employs the following procedure. A given ordering of the 12 tones within the octave is adopted. This tone row is repeatedly presented throughout the piece; however, the above transformations are allowed on each presentation. It is assumed that the row as an abstraction is perceived in its different manifestations.

Whether such transformations indeed result in perceptual equivalences is debatable. In the visual case we must have evolved perceptual mechanisms that preserve the identities of objects regardless of their orientation relation to the observer. An analogous argument cannot be made for the case of inversion and retrogression of sound sequences. A second doubt is based on general experience. Sound sequences may be unrecognizable when reversed in time, as the reader can determine by attempting to decode a segment of speech played backward. Furthermore, many inverted three-note combinations are perceptually very dissimilar. For example, the major and minor

[4]The use of the term "inversion" as defined by contemporary music theorists should not be confused with the traditional use of the term defined earlier.

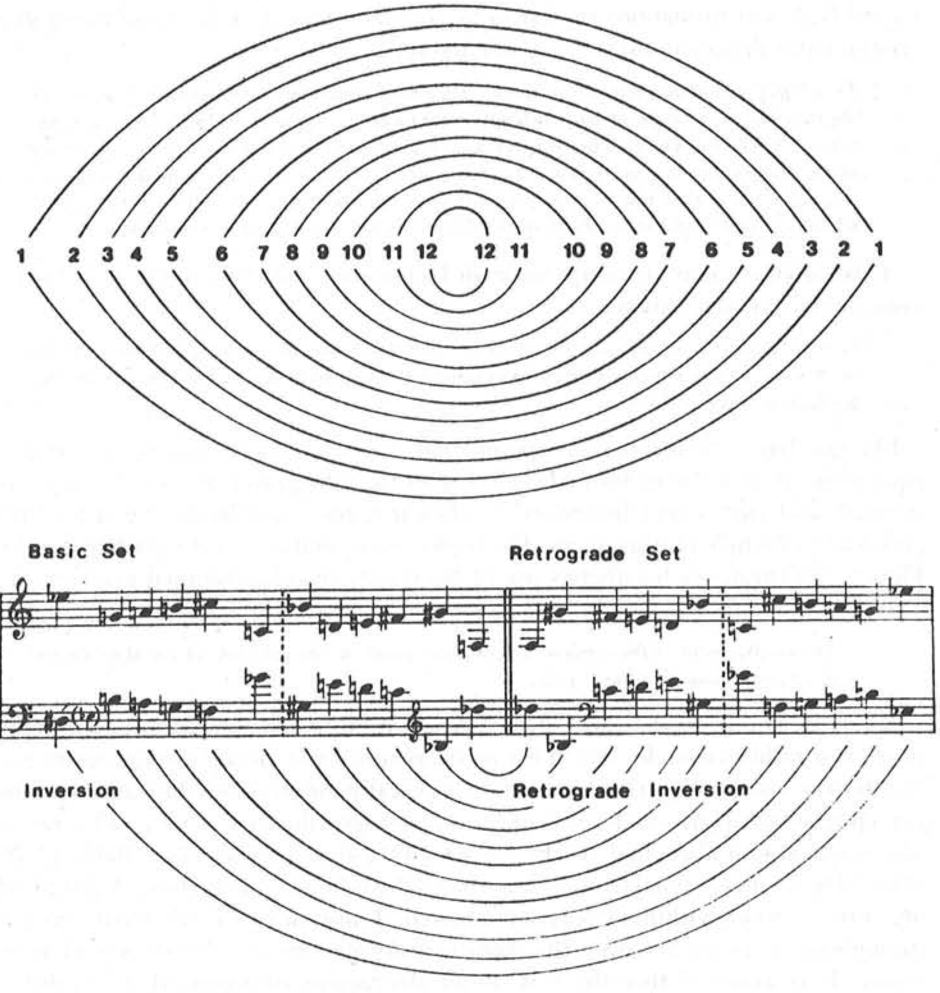

Fig. 7. Schoenberg's illustration of his theory of equivalence relations between pitch structures—taken from his *Wind Quartet, Op. 26* (from Schoenberg, 1951).

triads are inversions of each other; however, it appears quite implausible to regard them as perceptually equivalent. Rather it would seem that recognition of inverted and retrograde patterns is generally accomplished at a level of abstract coding equivalent to that which allows us to recite a segment of the alphabet backward, or to invert a sequence of numbers (Deutsch and Feroe, 1981).

At all events, a substantial body of music theory, aimed at defining equivalence and similarity relations between sets of pitches, is based on these assumptions of equivalence under retrogression and inversion, as well as the assumptions of pitch class and interval class identity. Although these theories are essentially concerned with stimulus structure, they do have implications about psychological representation.

A good example of this type of theory was provided by Chrisman (1971). He first defined the term "pitch-set" to refer to the set of elements derived from any collection of pitches, such that all members of a pitch class in the collection are represented by a unique element corresponding to that pitch class. In order to determine the intervallic relationships in a pitch-set, the elements of the pitch-set are given in ascending order, beginning with any pitch in the set. All members of the pitch-set are then contained within the octave above this initial pitch.

From any pitch-set whose elements have been placed in ascending order and within a single octave, one can construct an array which describes the intervallic structure of the set. Such an array consists of a linear succession of intervals and is termed a "successive-interval array." Each element of the array corresponds to the number of semitones between successive pitch-class representatives in the pitch-set. The array also includes the interval between the last note in the pitch-set and the first note an octave higher.

Cyclic reorderings of elements in each successive interval array are useful for determining intervallic relationships between pitch-sets with differing pitch contents. To take Chrisman's example, the pitch-sets S, T and V have different pitch contents and produce three different interval arrays: A, B, and C.

Pitch-Set	Interval-Array
S = C, C♯, E, F♯, G, A♯, B	A = 1-3-2-1-3-1-1
T = C, D♯, E, F, F♯, A, B	B = 3-1-1-1-3-2-1
V = C, D, D♯, F♯, G, G♯, A	C = 2-1-3-1-1-1-3

When the elements in these arrays are cyclically reordered, the arrays are shown to be equivalent.

$$
\begin{array}{rr}
 & A = 1\text{-}3\text{-}2\text{-}1\text{-}3\text{-}1\text{-}1 \\
B = 3\text{-}1\text{-}1\text{-}1\text{-}3\text{-}2\text{-}1 & P_3(\underline{B}) = 1\text{-}3\text{-}2\text{-}1\text{-}3\text{-}1\text{-}1 \\
C = 2\text{-}1\text{-}3\text{-}1\text{-}1\text{-}1\text{-}3 & P_5(\underline{C}) = 1\text{-}3\text{-}2\text{-}1\text{-}3\text{-}1\text{-}1
\end{array}
$$

The sets T and V are, thus by this definition, shown to be transpositions of the set S.

Another, much more elaborate formulation was proposed by Forte (1973). Forte was concerned not only with the conditions under which sets of pitches should be considered equivalent, but also with defining similarity relationships between pitch-sets. He used two measures of similarity, one based on pitch class intersection and the other on interval class intersection. For other work on pitch-sets see Howe (1965), Lewin (1960, 1962), Perle (1972, 1977), and Teitelbaum (1965).

The extent to which the structures defined by such theories are processed by the listener remains to be determined. As noted by Garner (1974), some structures that exist in a stimulus configuration are perceived readily, others with difficulty, and yet others not at all. A fundamental problem with this body of theory concerns the basic equivalence assumptions on which it rests. The issue of interval class is a thorny one, and the assumptions of equivalence under retrogression and inversion are also debatable. Reservations about these equivalence assumptions have also been raised recently by music theorists (Browne, 1974; Howe, 1974; Benjamin, 1974).

Other theorists have attempted to represent pitch relationships in terms of distances in a multidimensional space. Drobisch (1846, 1855) proposed that pitch be represented in three dimensions as a helix, with the vertical axis corresponding to pitch height, and tones separated by octaves lying closest within each turn of the helix. This representation reflects the perceptual closeness of the octave relationship. Shepard (Chapter 11) provides a detailed theoretical formulation that elaborates on this model.

Longuet-Higgins (1962a,b) has proposed that "tonal space" be represented as a three-dimensional array. Tones 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 separated by octaves are treated as equivalent, an array is obtained such as shown on Fig. 8. Note that a closely related set of tones, such as comprise the C-major scale, forms a compact group, so that a key can be defined as a neighborhood in this space. Longuet-Higgins proposed that when presented with a segment of music, the listener initially selects a given region of space, thus attributing a key. However, if his choice forces him to engage in large jumps

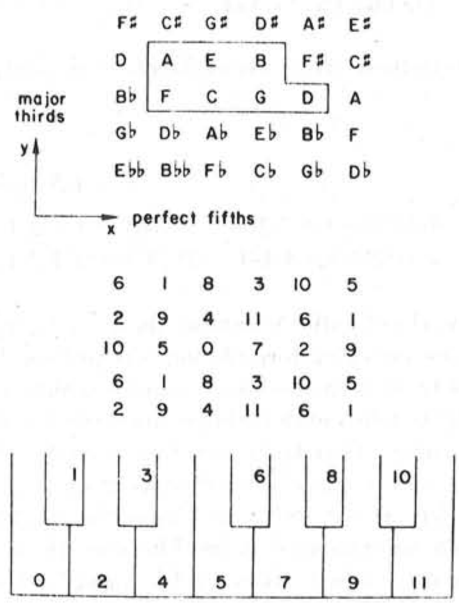

Fig. 8. Array hypothesized by Longuet-Higgins for the representation of "tonal space." See text for details (from Longuet-Higgins, 1978).

within this region, the listener abandons it and selects instead a region where the tones are more compactly represented, thus attributing a new key.

Another approach to the mapping of tonal space was taken by Krumhansl (1979). She performed an experiment in which subjects were presented with a set of context tones which were followed by two tones played in succession. The context tones were either the chord of the C-major triad or the C-major scale. Subjects judged on each trial how similar the first tone was to the second in the tonal system suggested by this context. Multidimensional scaling of the similarity ratings produced a three-dimensional conical structure around which tones were ordered according to pitch height. The components of the major triad formed a closely related cluster near the vertex of the cone; the remaining tones of the diatonic scale formed a less closely related subset that was further from the vertex; and the nondiatonic tones were widely dispersed, still further from the vertex. It is expected that different contexts would give rise to different patterns of similarity relationships using this technique. In the traditional music of our culture, the minor scale should produce a different configuration, and entirely different patterns would be expected from those familiar with other types of music. However, the study is important in demonstrating that pitch relationships are represented in a complex and well-defined fashion in a highly overlearned tonal context.

IV. ALPHABETS AND HIERARCHIES

We next consider a further level of abstraction, in which pitch information is mapped onto a relatively small set of highly overlearned alphabets. Although these differ from one culture to another, the use of such alphabets appears to occur cross-culturally. The invocation of a relatively small number of alphabets, each of which consists of a relatively small number of steps, allows for music of considerable complexity without the penalty of heavy processing load (Miller, 1956; Garner, 1974). In the tonal music of our tradition, the 12-tone chromatic scale forms a parent alphabet from which a family of subalphabets is derived, such as major scale, minor scale, major triad, and so on. Each of these subalphabets is itself a family of subalphabets that are related by transposition. Other alphabets involve chord progressions, such as progression along the cycle of fifths.

The ready mapping of musical sequences onto pitch alphabets is reflected in the finding that short-term transposition often tends to occur along such alphabets rather than in terms of exact interval sizes (Fig. 9). The consequent alterations in interval size do not produce an impression of musical "incorrectness." This contrasts with transposition in long-term situations, where exact intervals are generally preserved instead (Deutsch, 1969; Attneave & Olson, 1971). Short-term perceptual equivalences, therefore, appear to be more heavily influenced by pitch content than are long-term perceptual equivalences.

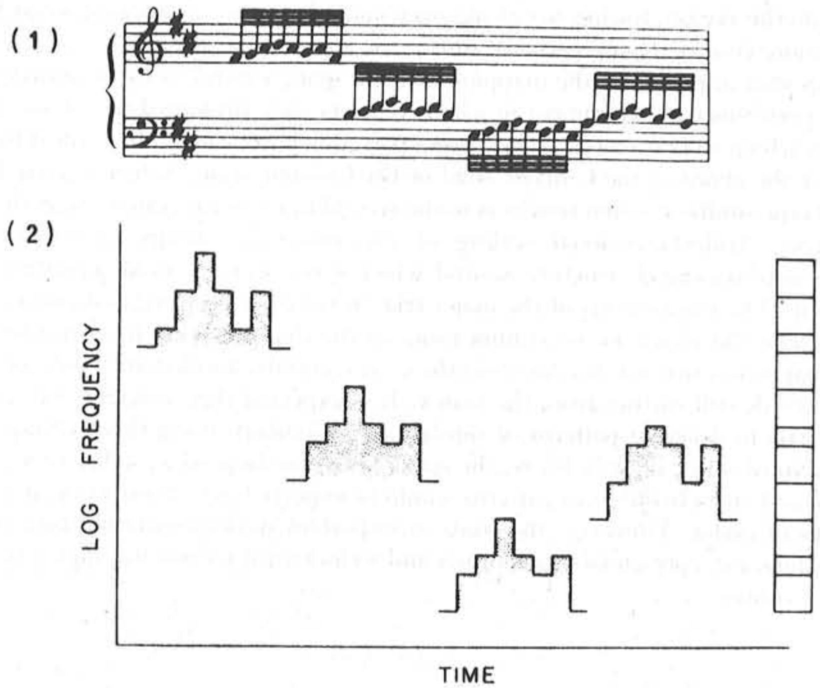

Fig. 9. Transposition along the alphabet of the scale. A given configuration is presented four times in succession at different positions along the scale. Since the scale has unequal intervals, there results a variation in the set of intervals involved. Ladder at right displays scale. From Bach, J. S. *The Well-Tempered Clavier, Book I, Fugue V* (from Deutsch, 1977).

The invocation of scalar alphabets in short-term situations has been studied by several investigators. Francès (1958) found that listeners were better able to detect alterations in tonal than in atonal melodies. Similar conclusions were reached by Dewar (1974) and Dewar, Cuddy, and Mewhort (1977). Dewar also observed that for tonal melodies, when the altered tone was in the same diatonic scale as the other tones, discrimination accuracy was poorer than when the altered tone departed from this scale. Dowling (1978) further observed that listeners had considerable difficulty in distinguishing between exact transpositions of melodies to new keys and shifts along the same diatonic scale where intervals were not preserved.

The tonal music of our tradition is also composed of small segments that are systematically organized in hierarchical fashion. It is reasonable to suppose that such hierarchical organization reflects the ways in which musical information is abstracted and retained. As Greeno and Simon (1974) point out, many different types of information appear to be retained as hierarchies. In some instances, the information stored is in the form of concepts that refer to classes (Collins & Quillian, 1972). We also appear to retain hierarchies of rules (Gagné, 1962; Scandura, 1970), hierarchies of programs (Miller, Galanter, & Pribram, 1960), and hierarchies of goals in problem solving (Ernst & Newell, 1969). Visual scenes appear to be represented as hierarchies of subscenes (Winston, 1973; Palmer, 1975; Navon, 1977; Hanson & Riseman, 1978). The phrase structure of a sentence lends itself readily to hierarchical interpretations (Yngve, 1960; Miller & Chomsky, 1963).

Experiments by Restle and Brown (1970 and Restle (1970) have demonstrated that we readily acquire serial patterns as hierarchies that reflect the structures of these patterns. In their experiments, subjects were presented with a row of six lights, which came on and off in repetitive sequence, and their task was to predict which light would come on next. The sequences were structured as hierarchies of operators. For example, if the basic subsequence is X = (1 2), then the operation R ("repeat of X") produces the sequence 1 2 1 2; the operation M ("mirror-image of X") produces 1 2 6 5; and the operation T ("transposition +1 of X") produces 1 2 2 3. By recursive application of such operations, long sequences can be generated that have compact structural descriptions. For example, the sequence 1 2 1 2 2 3 2 3 6 5 6 5 5 4 5 4 can be described as M(T(R(T(1)))) and corresponds to the *structural tree* shown on Fig. 10.

Using sequences constructed in this fashion, Restle and Brown (1970) showed that the probability of error in prediction increased monotonically with the level of transformation along such a structural tree. For example, the highest probability of error in a sequence such as on Fig. 10 occurred at Locations 1 and 9, the next highest at Locations 5 and 13, and so on. From this and other evidence, it was concluded that the observer organizes information in accordance with such structures. Parallel theoretical developments by Simon and his colleagues (Simon & Kotovsky, 1963; Simon & Sumner, 1968; Simon, 1972) and by others (Leewenberg, 1971; Jones, 1974, 1978; Vitz & Todd, 1967, 1969) also utilized hierarchies of operators.

Deutsch and Feroe (1981) argue that sequences in tonal music have characteristics not reflected in the above formalisms, and they advance a model for the encoding of pitch sequences as hierarchies of operators, which takes these characteristics into account. In essence, the model may 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 one level are elaborated by further elements so as to form structural units at the next-lower level, until the lowest level is reached. It is also

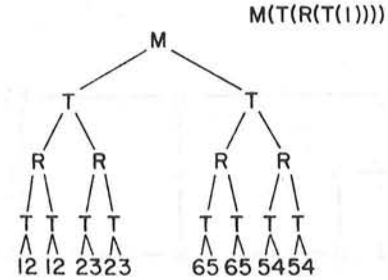

Fig. 10. Tree diagram of a long, regular binary pattern (from Restle, 1970).

assumed that Gestalt principles such as Proximity and Good Continuation contribute to organization at each hierarchical level.

The following are the rules for a simplified version of the system; however, the reader is referred to Deutsch and Feroe (1981) for a description of the full system:

1. A *structure* is notated by $(A_1, A_2, \ldots A_{l-2}, A_{l-1}, *, A_{l+1}, A_{l+2}, \ldots, A_n)$, 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 $(A_1, A_2, \ldots, *, \ldots, A_n)$ 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 $(A_1, A_2, \ldots, *, \ldots, A_n)$, 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 $A = (A_1, A_2, \ldots, *, \ldots, A_m)$, α, and $B = (B_1, B_2, \ldots, *, \ldots, B_n)$, β, where α and β are two alphabets, we define the compound operator pr (prime). $A|\text{pr}|B$;r, where r is the reference element, refers to assigning values to the notes produced from $(B_1, B_2, \ldots, *, \ldots, B_n)$, 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 $(B_1, B_2, \ldots, *, \ldots, B_n)$, 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.

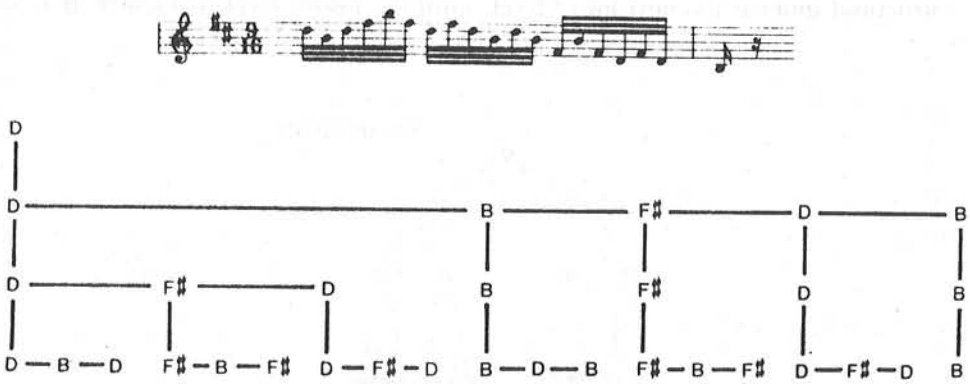

Fig. 11. Example to illustrate the model of Deutsch and Feroe for the internal representation of pitch sequences (see text for details). Passage is from Bach's *Sinfonia 15*, BWV 801 (from Deutsch & Feroe, 1981).

To give a simple example of the use of the formalism, Sequence 1 of Fig. 16 is represented as

$$A = (*,3p)G_{tr}$$
$$B = (*,p,n)Cr$$
$$S = A[pr]B,G_5$$

where G_{tr} represents the G-major triad, Cr the chromatic scale, and G_5 the reference element.

It may be observed that this form of representation results in considerable parsimony of encoding. Further, the employment of distinct alphabets at different structural levels enables an encoding in terms of proximal relationships, so that the principle of proximity is upheld (see Chapter 4).

A more complex musical example is shown on Fig. 11 and is represented as on three structural levels as follows:

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

So far we have considered 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 often implied. It is assumed that such progressions may also be encoded in hierarchical fashion. 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, and this in turn serves to reduce processing load.

The present model may be related to characterizations of hierarchical structure in tonal music advanced by music theorists. The most important development in this field is that of Heinrich Schenker (1868–1935), who proposed a hierarchical system for tonal music that has points of similarity with the system proposed by Chomsky for linguistics (Chomsky, 1963). On Schenker's system music is considered as a hierarchy in which pitch events at any one level are considered "prolonged" by sequences of pitch events at the next-lower level. Three basic levels are distinguished in the system. First, there is the surface representation or *foreground*; second, there is the *middleground*; and third, there is the *background* or *Ursatz*. The *Ursatz* is itself considered a prolongation of the triad. For literature on Schenkerian analysis and related theoretical formulations, see particularly Lerdahl and Jackendorff (1977), Meyer (1973), Narmour (1977), Salzer (1962), and Schenker (1956,1973) and Yeston (1977).

V. MEMORY SYSTEMS

We assume that separate memory systems exist for retaining information at different levels of abstraction. Craik and Lockhart (1972) have made the general argument that the higher the level of abstraction, or "depth of processing" of information, the longer its persistence in memory. This may indeed be true of music. It is clear from general experience that memory for melodic and harmonic intervals persists for considerably longer than memory for absolute pitch values (Deutsch, 1969; Attneave & Olson, 1971); memory for higher order abstractions may well persist for longer still. Such differences in the persistence of memory would have the consequence that when retention of a musical sequence is examined following different time periods, it is likely to reflect different forms of encoding.

In previous sections we have examined the question of encoding in detail, and we now turn to a consideration of the influences acting on information in storage. This question has been explored in depth for the system retaining absolute pitch values, and a little is also known about the system retaining interval information, and the system retaining higher level abstractions.

A. The System Retaining Absolute Pitch Information

When listeners make pitch comparison judgments between tones that are separated by a silent retention interval, accuracy declines gradually as the retention interval is lengthened (Koester, 1945; Harris, 1952; Bachem, 1954; Wickelgren, 1966, 1969). However, the rate of memory deterioration here is very slow. For example, Harris (1952) found that following a retention interval of 15 sec, a difference of a small fraction of a semitone could still be reliably discriminated. This stands in sharp contrast to comparison performance when a sequence of extra tones is interpolated during the retention interval. In an experiment by Deutsch (1970a), listeners were selected for obtaining a score of 100% correct in comparing tone pairs that were separated by a six-second retention interval. The test tones were either identical in pitch or they differed by a semitone. When eight extra tones were then interpolated during this interval, the error rate rose to 40%, even though the listeners were instructed to ignore the interpolated tones. In a further experiment by Deutsch (1970b), this performance decrement was shown not to be due to attention distraction nor to a general memory overload, since only a minimal decrement occurred when spoken numbers were interpolated instead. This was true even when the subjects were simultaneously required to recall the interpolated numbers. It was concluded that pitch memory is subject to interference caused specifically by other tones.

Further experiments demonstrated that tones interact with each other in memory in an orderly and systematic fashion. Such interactions were shown to be a function of both the pitch relationships between the interacting tones and their closeness in serial position.

In one experiment, the effect of a tone that formed part of a sequence interpolated between two test tones was studied as a function of its pitch relationship to the first test tone (Deutsch, 1972b). Subjects compared the pitches of two tones that were separated by a five-second retention interval during which a sequence of six extra tones was interpolated. The test tones were taken from the 12-tone chromatic scale, and ranged from Middle C to the B above. In half of the sequences the test tones were identical in pitch; in the other half they differed by a semitone. The pitches of the intervening tones were also taken from the 12-tone chromatic scale and ranged from the F♯ below Middle C to the F an octave and a half above. The tones were chosen at random from this range with certain restrictions.[5]

The experiment consisted of eight conditions. In all conditions but the last, a tone whose pitch bore a critical relationship to the pitch of the first test tone was placed in the second serial position of the intervening sequence. The relationship between the critical intervening tone and the first test tone varied between identity to a whole-tone separation. A unique value of pitch separation was incorporated in each of the seven conditions, these values being placed at equal intervals of 1/6 tone within this whole tone range. However, in the eighth condition, the pitch of the tone in the second serial position was chosen in the same way as were the other pitches in the intervening sequence. This last ("null") condition therefore provided a baseline against which the effects of the critical interpolated tone could be evaluated.

It was found that the effect of the critical interpolated tone on memory for the test tone varied systematically as a function of their pitch relationship. When the critical interpolated tone was identical in pitch to the first test tone, memory facilitation occurred. With increasing pitch separation between the first test tone and the critical interpolated tone, the error rate rose progressively; it peaked at a separation of 2/3 tone and declined to baseline at roughly a whole tone separation.

Further experiments have reinforced and elaborated on these findings of specific interactions within the pitch memory system. One study investigated the effects of including in an intervening sequence a tone that was a semitone removed from the first test tone (Deutsch, 1973b). It was found that when the test tones were identical in pitch, including in the intervening sequence a tone that was either a semitone higher or a semitone lower, an increase in errors resulted. When both a tone a semitone higher and also a tone a semitone lower were included in the intervening sequence, a substantially greater increase in errors occurred than when only one of these was included. Further, when the test tones differed in pitch by a semitone, including in the intervening sequence a tone that was identical in pitch to the second test tone; there was a substantial increase in errors. An increase in errors that was however significantly smaller also occurred when the critical intervening tone was a semitone removed from the first test tone, but on the opposite side of the pitch continuum to the second test tone. When both of these tones were included in the same intervening sequence, a significantly greater increase in errors was produced than when only one of these was included.

This experiment demonstrated at least two separable disruptive effects in pitch memory. First, the inclusion of a tone that was a semitone removed from the tone to be remembered produced a small but significant disruptive effect which cumulated in size when two such tones, one higher and the other lower than this tone, were both included. Second, a significantly larger disruptive effect occurred when the test tones

[5]No sequences contained two tones of the same pitch class. Further, all tones were excluded from any sequence that lay within, and including, a whole tone range in either direction from the first test tone, or that were displaced by an octave from this range. This gap was necessary to prevent the random inclusion of tones in the critical range under study.

differed in pitch and the critical interpolated tone was identical in pitch to the second test tone.

To explain the second (and larger) of these two disruptive effects, the following hypothesis was advanced (Deutsch, 1972c). Memory for the pitch of a tone is laid down simultaneously both on a pitch continuum and also on a temporal or order continuum. As time proceeds, this memory distribution spreads in both directions, but particularly along the temporal continuum. As a result of this spread, when a tone of the same pitch as the second test tone is presented in the interpolated sequence, the subject sometimes concludes that this had been the first test tone. That is, errors of misrecognition result from the subject correctly recognizing that a tone of identical pitch to the second test tone had occurred, but being uncertain when it had occurred.

This hypothesis, which is described in detail in Deutsch (1972c), gives rise to various predictions. First, we should expect errors of misrecognition to be more numerous when the critical interpolated tone is placed early in the intervening sequence rather than late, since here there would be a greater chance of temporal or order confusion. That is, the closer the critical interpolated tone is to the first test tone, the more difficult it should be to discriminate their two positions along a temporal or order continuum and so the greater the number of misrecognition errors.

So, in a further experiment subjects made pitch recognition judgments between tone pairs that were separated by a sequence of six interpolated tones. In sequences where the test tones differed in pitch, a tone that was identical in pitch to the second test tone was sometimes placed in the second serial position of the intervening sequence, sometimes in the fifth serial position, and sometimes no such tone was included. It was found that the critical interpolated tone produced an increase in errors at both serial positions compared with sequences containing no such tone; however, the effect was much more pronounced when the critical tone was placed early in the intervening sequence rather than late (Deutsch, 1975b). This finding therefore corroborated the hypothesis.

Also in this experiment, in sequences where the test tones were identical in pitch, the effect was studied of including a tone that was a semitone removed from the test tone, as a function of its serial position in the interpolated sequence. No serial position effect was here found. So it seems that the first source of disruption, in contrast, does not depend on the serial position of the critical interpolated tone.

Given this large serial position effect based on interpolating a tone of identical pitch to the second test tone, we can ask whether this is due to a deterioration of information along a temporal continuum, or whether this effect is a function of order rather than time. An experiment was performed to investigate this question. Subjects compared the pitches of two tones that were separated by a sequence of six interpolated tones. In sequences where the test tones differed in pitch, a tone identical to the second test tone was placed either in the second serial position of the intervening sequence, or in the fifth serial position, or no such tone was interpolated. The experiment studied the effect of lengthening the pause following the first test tone, and also of lengthening the pause preceding the second test tone. The different temporal conditions are shown on Fig. 12. Now if this serial position effect were due simply to

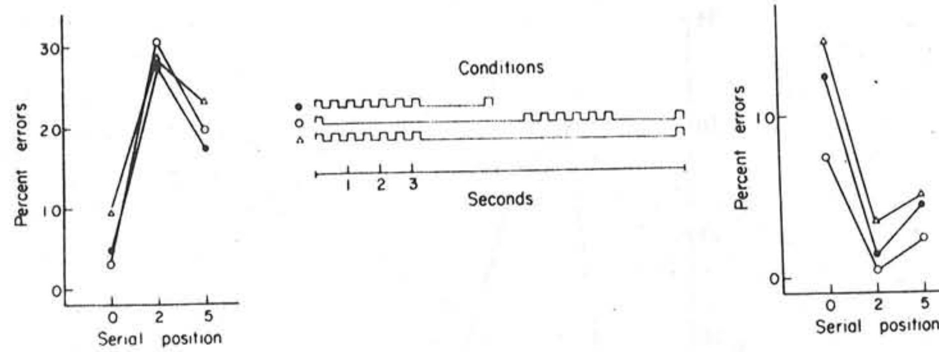

Fig. 12. Temporal parameters used in two experiments investigating the effects of unfilled delays on pitch recognition. Graph on left plots error rates in sequences where the test tones differed and a tone whose pitch was identical to the second test tone was placed in the intervening sequence. Graph on right plots error rates in sequences where a tone of identical pitch to the first test tone was interpolated. Numbers 2 and 5 indicate that the critical tone was placed in the second and fifth serial positions of the intervening sequence, respectively; "0" indicates that no such tone was interpolated.

interactions along a temporal continuum, we should expect substantial changes in the shape of the effect to result from such temporal manipulations. However, as shown on Fig. 14, the serial effect occurred under all temporal conditions. It would appear, therefore, that this deterioration of information takes place along an order continuum, which is not sensitive to temporal variations, at least in the range investigated here (Deutsch, in preparation).

There is a second prediction from the hypothesis of loss of temporal or order information. Suppose that we plotted errors precisely as a function of the pitch relationship between the first test tone and a critical interpolated tone, in 1/6 tone steps as described earlier, and we also varied the pitch difference between the two test tones. Then, in sequences where the critical interpolated tone and the second test tone are placed on the same side of the pitch continuum relative to the first test tone, the peak of errors should occur where the critical interpolated tone is identical in pitch to the second test tone. That is, a shift in the pitch of the second test tone, when the pitch of the first test tone is held constant, should result in a parallel shift in the peak of errors produced by the critical interpolated tone.

This prediction was tested in another experiment (Deutsch, 1975b). Here, when the test tones differed in pitch, this difference was either 1/3 tone, 1/2 tone, or 2/3 tone. Errors were plotted as a function of the pitch of a tone placed in the second serial position of an interpolated sequence, whose relationship to the pitch of the first test tone varied in 1/6 tone steps between identity and a whole tone separation. Whenever the second test tone was higher in pitch than the first, the critical interpolated tone was also higher; whenever the second test tone was lower, the critical interpolated tone was also lower. So, when the second test tone and the critical interpolated tone were separated from the first test tone by the same pitch distance, they were also identical in pitch to each other.

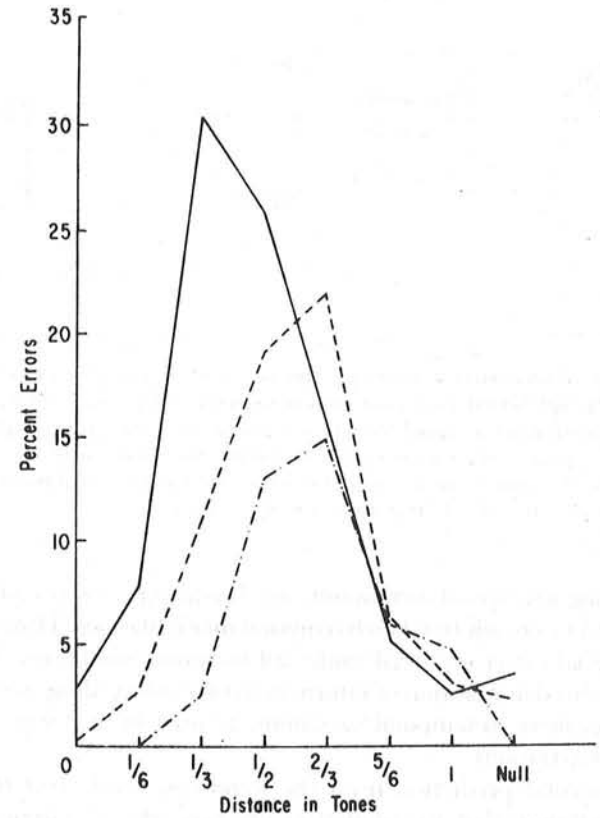

Fig. 13. Percent errors in pitch comparisons in sequences where the test tones differed in pitch, and the critical interpolated tone was placed on the same side of the first test tone along the pitch continuum. Errors were plotted as a function of the pitch relationship between the first test tone and the critical interpolated tone. (—) Test tones separated by 1/3 tone. (---) test tones separated by 1/2 tone. (---) test tones separated by 2/3 tone (from Deutsch, 1975b).

The results of this experiment are shown in Fig. 13. It can be seen that when the test tones were 1/3 tone apart, errors peaked when the critical interpolated tone was 1/3 tone removed from the first test tone, and so when it was identical in pitch to the second test tone. Similarly, when the test tones were 2/3 tone apart, errors peaked when the critical interpolated tone was 2/3 tone removed from the first test tone, and so again when it was identical in pitch to the second test tone. So, here a shift in the pitch of the second test tone relative to the first did indeed result in the predicted shift in the peak of errors produced by the critical interpolated tone. But note also that when the first and second test tones were a semitone apart, errors peaked not at a semitone but at 2/3 tone. This can be explained by assuming that the first source of disruption that peaks at 2/3 tone is superimposed on the present source of disruption produced by a relationship of identity or near identity between the second test tone and the critical interpolated tone.

In considering possible bases for this first source of disruption, two points should be noted. The first is that the relative frequency range over which this 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, 1969, 1970). Second, we can remember that the error rate cumulates when two such disruptive tones are interpolated, placed one on either side of the test tone along the pitch continuum (Deutsch, 1973b). There is an analogy here with lateral inhibitory interactions found in systems handling sensory information at the incoming level, where there is a cumulation of inhibition from stimuli placed on either side of the test stimulus (Ratliff, 1965). Evidence for lateral inhibition has been found in the system that handles pitch information at the incoming level (Carterette, Friedman, & Lovell, 1969, 1970; Houtgast, 1972). It was therefore theorized that elements of the pitch memory system are arranged as a recurrent lateral inhibitory network, analogous to those found in systems handling incoming sensory information. Elements of this system are activated by tones of specific pitch, and are organized tonotopically on a log frequency continuum (Deutsch & Feroe, 1975.)

Now, if this were so, we might hope to obtain an effect which would not be expected on other grounds. It has been found in physiological studies of peripheral receptors that when a unit that is inhibiting a neighboring unit is itself inhibited by a third unit, this releases the originally inhibited unit from inhibition. This phenomenon is known as disinhibition. Applying this to our present situation, we might expect that if a tone that was inhibiting memory for another tone were itself inhibited by a third tone, this could cause memory for the first tone to return. More specifically, in sequences where the test tones are identical in pitch, if two critical tones were interpolated—one always 2/3 tone removed from the first test tone and the other further removed along the pitch continuum—then the error rate should vary systematically as a function of the pitch relationship between the two critical interpolated 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. The curve produced should therefore be roughly the inverse of the curve plotting the original disruptive effect.

Accordingly, the following experiment was performed. Subjects compared the pitches of two test tones, which were separated by a sequence of six interpolated tones. A tone which was 2/3 tone removed from the first test tone was always placed in the second serial position of the interpolated sequence. Errors were then plotted as a function of the pitch of a further tone, which was placed in the fourth serial position, whose relationship to the tone in the second serial position varied in 1/6 tone steps between identity and a whole-tone separation. As can be seen in Fig. 14, a systematic return of memory was indeed obtained. The error rate in sequences where the second critical interpolated tone was identical in pitch to the first was significantly higher than baseline; further, the error rate where the two critical interpolated tones were separated by 2/3 tone was significantly lower than baseline. A first-order inhibitory function was also obtained experimentally, using subjects selected on the same criterion as for the disinhibition study. This function was then used to calculate the

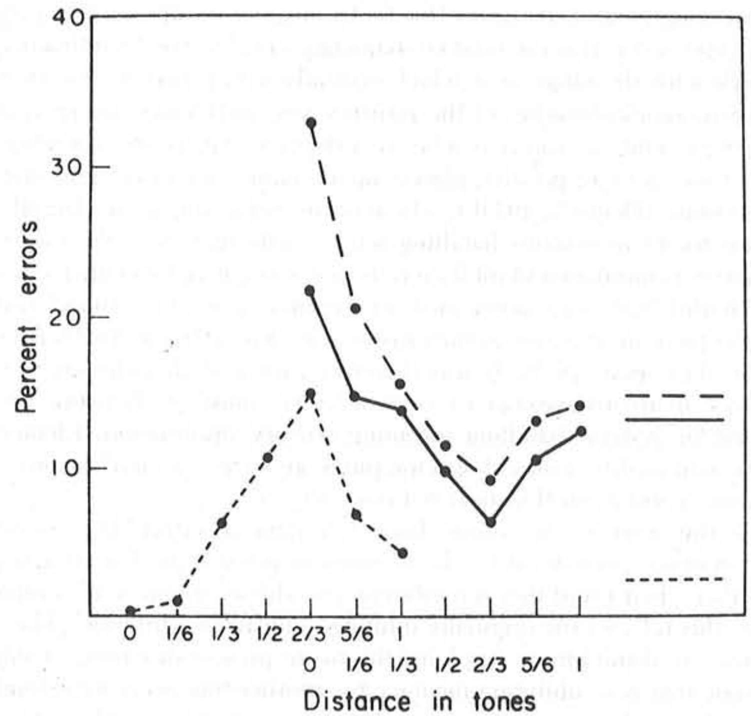

Fig. 14. Percent errors in pitch recognition obtained experimentally and predicted theoretically. Open triangles display percent errors in a baseline experiment that varied the pitch relationship between a test tone and a critical interpolated tone. (The open triangle at the right displays percent errors where no tones were interpolated in the critical range under study.) Filled circles display percent errors in an experiment where a tone that was 2/3 tone removed from the 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. Open circles display percent errors for the same experimental conditions predicted theoretically from the lateral inhibition model. (Filled and open circles at the right display percent errors obtained experimentally and assumed theoretically where no further critical tone was interpolated (from Deutsch & Feroe, 1975).

theoretical disinhibition function. As shown in Fig. 13, there is a good correspondence between the disinhibition function derived experimentally and that derived theoretically. This experiment provides strong evidence that pitch memory elements are indeed arranged as a lateral inhibitory network, analogous to those handling incoming sensory information.

1. Octave Generalization Effects

So far we have been considering interactions in memory between tones that are separated by less than an octave. We can now ask whether these effects take place along an array that is organized simply in terms of pitch, or whether an abstracted octave array, such as hypothesized on pp. 275–276, is also involved.

In one experiment (Deutsch, 1973a), subjects compared the pitches of two tones that were separated by a sequence of six interpolated tones. This experiment studied the effects of including in the interpolated sequence tones that bore the same relationships to the test tones as had been found earlier to produce disruption, but that were further displaced by an octave. In sequences where the test tones were identical in pitch, the effects were studied of interpolating two tones, one a semitone higher than the test tone and the other a semitone lower, except that the critical tones were also displaced an octave up or down. In sequences where the test tones differed, the effects were investigated of interpolating a tone of the same pitch as the second test tone, but that was also displaced an octave up or down. It was found that a substantial generalization of disruptive effect occurred from tones placed in the higher octave, and a weaker effect occurred from tones placed in the lower octave. However, the error rate was greatest when the critical tones were placed in the middle octave. From this pattern of results it was concluded that these disruptive effects take place along both a monotonic pitch continuum and also along an abstracted octave array.

2. *Memory Consolidation*

When a tone of identical pitch to the first test tone is included in an intervening sequence, the effect on memory is facilitatory rather than disruptive. In one experiment, subjects compared the pitches of two tones that were separated by a sequence of interpolated tones. In one condition, four tones were interpolated; in another, six tones were interpolated; in yet another, six tones were again interpolated, with a tone of identical pitch to the first test tone placed in the second serial position of the interpolated sequence. It was found that the error rate was lowest in sequences where the pitch of the first test tone was repeated, even compared with sequences containing fewer interpolated tones. We can conclude that the pitch memory system is subject to consolidation through repetition (Deutsch, 1975c).

However, in another experiment it was found that this consolidation effect is very sensitive to the serial position of the repeated tone. Here, subjects made pitch comparison judgments between tone pairs that were separated by a sequence of six interpolated tones. In some sequences a tone of the same pitch as the first test tone was placed in the second serial position of the interpolated sequence; in other sequences such a tone was placed in the fifth serial position; and in yet other sequences no such tone was interpolated. It was found that the facilitation effect was much more pronounced when the repeated tone was in the second serial position than in the fifth; indeed, this effect was statistically significant only for sequences where the repeated tone was in the second serial position (Deutsch, 1975c).

It was hypothesized that this consolidation effect results from the same process as produces the errors of misrecognition discussed above—namely, the spread of memory distribution along a temporal or order continuum. When two such distributions overlap, the overlapping portions sum, resulting in a stronger memory trace (Deutsch, 1972c). This again raises the question of whether this serial position effect is based on temporal factors, or whether it is simply a function of order. One might, for example, hypothesize that temporal proximity determines the amount of consolidation in pitch memory. If this were the case, then interpolating a long pause between the first test tone and the first interpolated tone should substantially reduce the consolidation effect. Alternatively, one might hypothesize that some mechanism tags the incoming stimuli in terms of order, and that consolidation takes place along such an order continuum, independent of temporal variations.

An experiment was therefore carried out to evaluate the effects of varying temporal parameters on the memory consolidation effect. Subjects compared the pitches of two tones that were separated by a sequence of six interpolated tones. A tone that was identical in pitch to the first test tone was placed either in the second serial position of the intervening sequence, or in the fifth serial position, or no such tone was included. The experiment studied the effect of lengthening the pause following the first test tone, and also of lengthening the pause preceding the second test tone. The different temporal conditions together with their error rates are shown on Fig. 12. These temporal manipulations did not significantly alter the consolidation effect as a function of serial position. It appears, therefore, that within this temporal range, consolidation takes place along a continuum that is organized in terms of order independent of time (Deutsch, in preparation).

B. The System Retaining Interval Information

We now briefly consider the influences acting on interval information in storage. Deutsch (1975d) proposed that memory for such information is based on a continuum whose elements are activated by the simultaneous or successive presentation of pairs of tones. Tone pairs that stand 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; and so on. It was further hypothesized that interactive effects take place along this continuum that are analogous to those that occur within the system retaining absolute pitch values. Such effects would include consolidation through repetition and similarity-based interference.

An experiment was performed to test this hypothesis (Deutsch, 1978b). Subjects compared the pitches of two test tones that 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. Thus, when the test tones were identical, the intervals formed by the test tone combinations were also identical. And 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 tone combination, the error rate was lower than when the sizes of the intervals formed by the interpolated combinations were chosen at random. Further, when the intervals formed by the interpolated combinations differed in size by a semitone from the interval formed by the first test tone combination, the error rate was higher than when the sizes of the intervals formed by the interpolated combinations were chosen at random.

This experiment, therefore, demonstrates the presence of both consolidation through repetition and also similarity-based interference in memory for harmonic intervals. This indicates that the system retaining such information is similar in organization to the system retaining absolute pitch values.

C. Interactions between These Systems

So far we have examined interactions that take place within a given memory system. We now turn to a consideration of how the outputs of the different systems interact in determining memory judgments. First, we examine ways in which abstracted information influences judgments of absolute pitch values. Then, we consider how information in the pitch memory system might in turn influence memory for pitch abstractions.

Since relational information is retained in parallel with pitch information, we might expect that judgments of sameness or difference in the pitches of two test tones would be biased by a sameness or difference in the relational context in which these test tones are placed. This possibility was investigated by Deutsch and Roll (1974). Subjects compared the pitches of two 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; and these patterns of relationship were present both when the test tones were identical in pitch and also when these differed. It was found that relational context had a strong influence on pitch recognition judgments. When the test tones were identical but were placed in different relational contexts, there resulted an increased tendency for their pitches to be judged as different. Further, when the test tones differed in pitch, but were placed in an identical relational context, there resulted an increased tendency for their pitches to be judged as identical. It was further found that when the test tones differed, and the test tone combinations formed intervals that were inversions of each other, there also resulted an increase in errors of misrecognition. It was concluded that this misrecognition effect was based on the perceptual equivalence of the inverted intervals.

Deutsch (in press) 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, analogous to that found for harmonic relational context.

Melodic relationships have been shown to influence pitch recognition judgments in another way. In listening to sequences, we process 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. Interpolated sequences forming melodic configurations that are more easily processed should then be associated with enhanced performance.

As described in Chapter 4, there is considerable evidence that melodic sequences are processed more effectively when these are composed of intervals of smaller size rather than larger (reflecting the operation of the principle of Proximity). One would therefore expect that interpolated sequences composed of smaller melodic intervals should be associated with higher performance levels than interpolated sequences composed of larger intervals. In an experiment by Deutsch (1978a), subjects compared the pitches of two tones that were separated by a sequence of six interpolated tones. Performance was studied under four conditions. In Condition 1 the interpolated tones were chosen at random from within a one-octave range, and they were ordered at random. Condition 2 was identical to Condition 1 except that the interpolated tones were arranged in monotonically ascending or descending order, with the result that the average size of the melodic intervals in the sequence was reduced. In Condition 3 the interpolated tones were chosen at random from within a two-octave range and were also ordered at random. Condition 4 was identical to Condition 3 except that the interpolated tones were arranged in monotonically ascending or descending order.

As shown on Fig. 15, the error rate was found to increase with an increase in the average size of the melodic intervals comprising the sequence. There was no evidence that monotonic ordering of the interpolated tones had any effect (apart from producing a smaller average interval size), though this might have been hypothesized from the principle of Good Continuation (see Chapter 4).

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 (Ortmann, 1926; Merriam, 1964; Fucks, 1962; Dowling, 1967; Jeffries, 1974; Deutsch, 1978d).

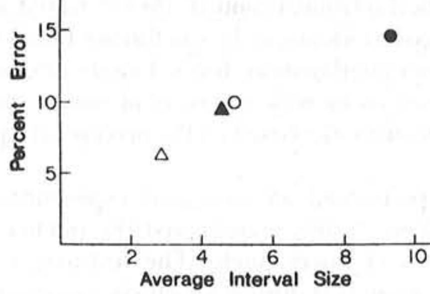

Fig. 15. Percent errors in pitch comparisons as a function of the average size of melodic interval in the sequence. Interpolated tones span a one-octave range: ▲ tones ordered at random; ● tones ordered monotonically. Interpolated tones span a two-octave range: ● tones ordered at random; ○ tones ordered monotonically (from Deutsch, 1978c).

One might hypothesize that this tendency is based on an increasing difficulty in the processing of melodic intervals as the sizes of these intervals increase. As indicated from the present experiment, this should in turn result in decreased accuracy in pitch recognition.

Two related studies should here be cited. Deutsch (1974) had subjects compare the pitches of two test tones that were separated by a sequence of eight interpolated tones. In one condition the interpolated tones were all drawn from the same octave as the test tones. In a second condition the interpolated tones were all drawn from the octave higher. In a third condition they were all drawn from the octave lower. In a fourth condition half of the interpolated tones were drawn from the octave higher than the test tones and the other half from the octave lower, the order of octave placement being random. It was found that when the interpolated tones were drawn from a single adjacent octave, the error rate was lower than when they were all drawn from the same octave. However, when the interpolated tones were drawn from both the octave higher and also the octave lower, the error rate was highest of all. In this last condition the average size of the intervals formed by successive tones in the sequence was considerably larger than in the other conditions, and it was concluded that these large jumps made it very difficult to make use of the melodic information in the interpolated sequence.

Another related study is that of Olson and Hanson (1977). They found, using several different paradigms, that increased errors in pitch recognition were associated with an increased pitch distance between the test tones and the interpolated tones. Since only three interpolated tones were used in this study, this increased pitch distance resulted in a significant increase in average interval size. These authors suggest an interpretation of their results that is very similar to the present line of reasoning.

Evidence for interactions based on more complex encodings has also been obtained. Krumhansl (1979) required subjects to compare the pitches of two test tones that were separated by a sequence of interpolated tones. Some of the sequences were constructed so as to suggest a tonality. It was found that for such sequences, when the first test tone was in the same scale as the interpolated tones, recognition performance was better than when it was outside this scale.

We should expect that interactions taking place within the pitch memory system would in turn affect memory for abstracted information. For example, given the memory consolidation effect produced by a repeated tone (Deutsch, 1975c), we should expect that sequences containing repeated tones would be better remembered, even in abstracted form, than sequences not containing repeated tones. Similarly, given the large disruptive effect produced by interpolating two tones, one a semitone higher than the tone to be remembered and the other a semitone lower (Deutsch, 1973b), we should expect that music composed in scales consisting mostly of semitonal steps would be more difficult to remember than music composed in scales consisting mostly of whole tone steps.

Erickson (1978) has argued that low-level effects, such as described here, have probably exerted a strong influence on the evolution of music systems. For example, it appears true cross-culturally that musical sequences tend to contain one or two tones that are repeated considerably more often than others. In our tonal system, for example, the first tone of the scale (the tonic) has this property. Given the finding of consolidation through repetition, we should expect that such repeated tones should be particularly well remembered. Second, it appears true cross-culturally that small melodic intervals occur considerably more often than large ones. We have also seen that recognition of the pitch of a tone is better in the context of a sequence consisting of small melodic intervals rather than of large ones (Deutsch, 1978c). When we consider these two low-order effects together, we can see that a considerable processing advantage is to be gained from a system in which there are a limited number of anchor tones—which are well remembered through repetition—surrounded by satellite tones which are linked to these anchor tones by pitch proximity. Thus, examination of these low-order effects enables us to develop an understanding of why musical systems have evolved as they did. Such systems should therefore not be considered as arbitrary sets of rules but rather as reflecting constraints imposed by our processing mechanisms.

D. Memory for Hierarchically Structured Tonal Sequences

We now turn to a consideration of memory for tonal information which is projected onto highly overlearned pitch alphabets, and which is organized in hierarchical fashion. Deutsch and Feroe (1981) have proposed that, at this level, tonal sequences are coded and retained as hierarchies of structures, each of which is associated with a given pitch alphabet. It is further proposed that such structures are encoded and retained as chunks, along with their associated alphabets and rules of combination. If this model is correct, then sequences that may be parsimoniously represented according to its rules should be better retained than those where such parsimonious representation is not possible.

An experiment was performed as a test of this hypothesis (Deutsch, 1980, 1981). In this experiment a second factor was also considered. It has been found in studies using strings of verbal materials that we tend to recall such strings in accordance with their temporal grouping (Mueller & Schumann, 1894; McLean & Gregg, 1967; Bower & Winzenz, 1969). This effect can be so powerful as to mask grouping by meaning, and so to obliterate the advantage incurred by such grouping (Bower & Springston, 1970). Analogous results have been obtained using nonverbal materials (Restle, 1972; Handel, 1973; Dowling, 1973). It was therefore expected that temporal grouping would have a strong effect on recall of these sequences also. Grouping in accordance with tonal structure was expected to result in enhanced performance, and grouping in conflict with tonal structure to give rise to performance decrements.

The experiment employed the following paradigm. Musically trained subjects were presented with sequences which they recalled in musical notation. The sequences employed are shown in Fig. 16. It can be seen that those in Fig. 16A each consisted of a higher level subsequence of four elements that acted on a lower level subsequence of

Fig. 16. Sequences used in first experiment to study utilization of structure in recall. See text for details (from Deutsch, 1981).

three elements. The sequences in Fig. 16B cannot be represented in this parsimonious fashion. These eight sequences were each presented in three temporal configurations, which are illustrated in Fig. 17. In the first configuration (a), the tones were spaced at equal intervals; in the second (b), they occurred in four groups of three, so that segmentation was in accordance with tonal structure; in the third (c), they occurred in three groups of four so that segmentation was in conflict with tonal structure.

Table I shows the percentages of tones correctly recalled in their correct serial positions in the different conditions of the experiment. Large effects of tonal structure and temporal segmentation are apparent. For structured sequences that were segmented in accordance with tonal structure, the performance level was extremely high. For structured sequences that were unsegmented, the performance level was slightly lower, though still very high. However, for structured sequences that were segmented in conflict with tonal structure, the performance level was much lower. For unstructured sequences the performance level was considerably lower than for structured sequences that were segmented in accordance with tonal structure or that were

Fig. 17. Types of temporal segmentation employed in first experiment to study utilization of structure in recall. a. Sequence unsegmented. b. Sequence segmented in groups of three, so that segmentation was in accordance with structure. c. Sequence segmented in groups of four, so that segmentation was in conflict with structure.

TABLE I

Utilization of Structure in Recall: Percent Correct Recall of Tones in Correct Serial Positions in Experiment 1[a]

Condition	Percent correct
Sequences structured in groups of 3	
OS. Not temporally segmented	93.5
3S. Temporally segmented in groups of 3	99.3
4S. Temporally segmented in groups of 4	69.2
Sequences unstructured	
OU. Not temporally segmented	52.0
3U. Temporally segmented in groups of 3	63.2
4U. Temporally segmented in groups of 4	62.3

[a] From Deutsch (1981).

unsegmented, but in the same range as for structured sequences that were segmented in conflict with tonal structure.

The serial position curves for the different conditions of the experiment are shown on Fig. 18. Typical bow-shaped curves are apparent, and in addition, discontinuities can be seen at the boundaries between temporal groups. This provides further evidence that temporal groups tend to be encoded as chunks and to be retained or lost independently. This type of configuration is very similar to that obtained by others with the use of verbal materials (Bower & Winzenz, 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 to be larger for the transition into the first element of a chunk. The TSP values for sequences segmented in temporal groups of three and four are shown in Fig. 19A and B, respectively. It can be seen that the TSPs are larger on the first element of each temporal group than on the other elements. This is as expected on the hypothesis that temporal groups serve to define subjective chunks that are retained or lost independently of each other.

Finally, a very strong sensitivity to musical alphabet was demonstrated in this experiment. As shown in Fig. 16, four sequences employed only a triadic alphabet, two employed a diatonic alphabet, and two employed other tones from the 12-tone chromatic scale. Of the 12 subjects who participated in the experiment, 6 remained entirely within the alphabet of the sequence they were notating. (For example, in notating sequences 2, 4, 6, and 8, all their incorrect responses, as well as their correct ones, were G, B, or D.) Of the remaining six subjects, five made between them a total of only 15 responses that departed from the alphabet presented. This implies that the subjects were retaining information concerning alphabet independently of structure.

Now in this experiment, the structured sequences all consisted of a higher level

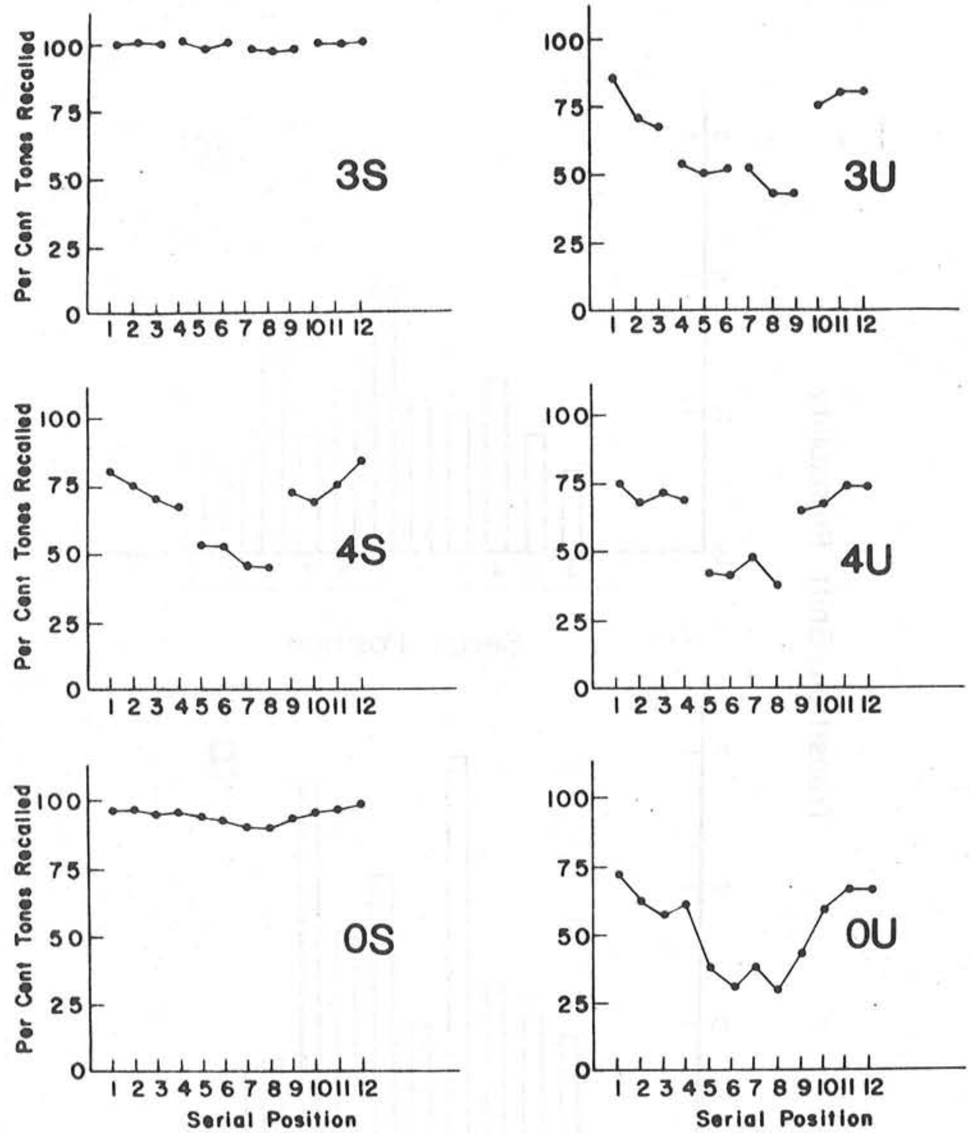

Fig. 18. Serial position curves for the different conditions of the first experiment to study utilization of structure in recall (from Deutsch, 1981).

subsequence of four elements that acted on a lower level subsequence of three elements. So compatible segmentation was always in groups of three and incompatible segmentation in groups of four. One might therefore argue that the enhanced performance found for structured sequences was due simply to an advantage conferred by the size of temporal group. A second experiment was designed to control for this.

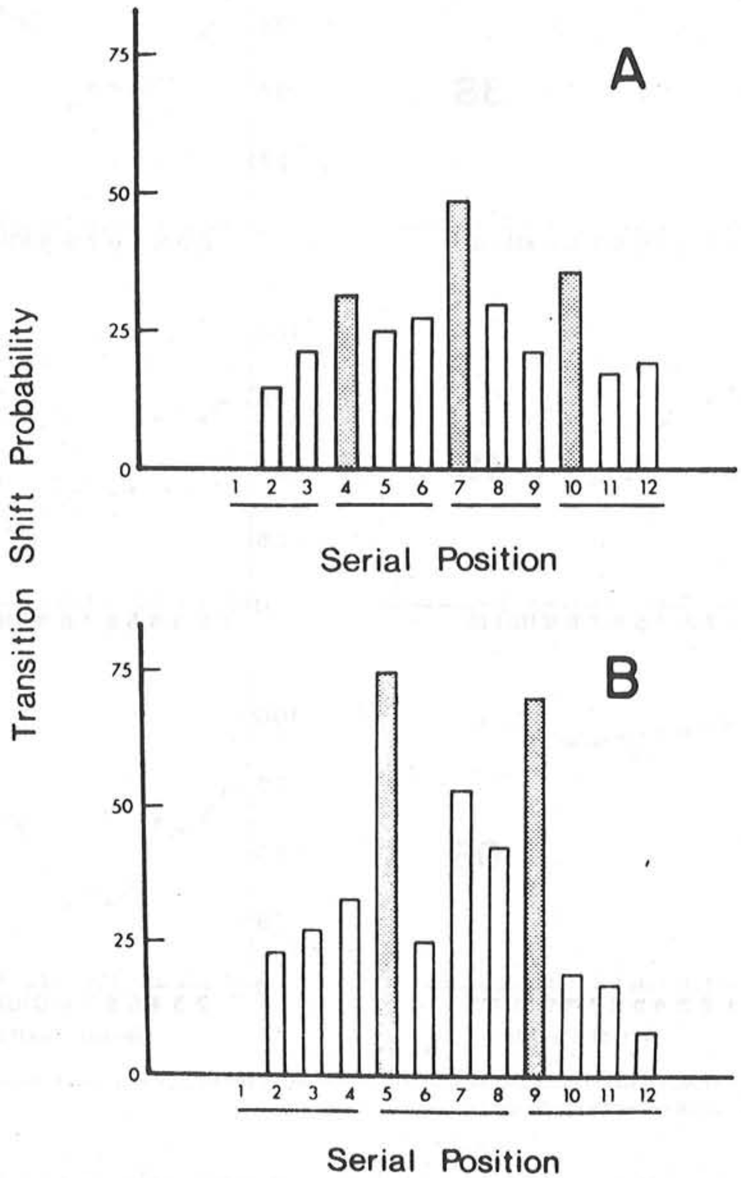

Fig. 19. (A) Transition shift probabilities for sequences segmented in temporal groups of three. (B) Transition shift probabilities for sequences segmented in temporal groups of four (from Deutsch, 1981).

Fig. 20. Sequences used in second experiment to study utilization of structure in recall. See text for details (from Deutsch, 1981).

Two types of tonal structure were employed, and these are shown in Fig. 20. In the first, shown in Fig. 20A, a higher level subsequence of four elements acted on a lower level subsequence of three elements. In the second, shown in Fig. 20B, a higher level subsequence of three elements acted on a lower level subsequence of four elements. Enhanced performance was expected where the number of elements in the lower level subsequence corresponded to the number within a temporal group, compared with sequences where these numbers did not correspond. In addition, segmentation in groups of two was examined. It was expected that where the lower level subsequence consisted of three elements, such segmentation would result in considerable performance decrements since it would conflict with tonal structure. However, where the lower level subsequence consisted of four elements, segmentation in group of two would be less disruptive since pauses would still be placed between tonal groups.

There were, therefore, six conditions in the experiment, shown in Table II. Also shown are the percentages of tones correctly recalled in their correct serial positions in these different conditions. It can be seen that performance levels were considerably lower when temporal segmentation was in conflict with structure. When pauses were

TABLE II
Utilization of Structure in Recall:
Percent Correct Recall of Tones in Correct Serial Positions in Experiment 2[a]

Condition	Percent correct
Sequences structured in groups of 3	
3-2. Temporally segmented in groups of 2	45.4
3-3. Temporally segmented in groups of 3	93.1
3-4. Temporally segmented in groups of 4	50.6
Sequences structured in groups of 4	
4-2. Temporally segmented in groups of 2	80.8
4-3. Temporally segmented in groups of 3	52.9
4-4. Temporally segmented in groups of 4	85.4

[a] From Deutsch (1981).

placed both between and within tonal groups (that is Condition 4-2), the performance level was slightly lower than when the pauses were placed only between groups, but considerably higher than when the pauses conflicted with tonal structure. The interaction between size of tonal unit and size of temporal unit was highly significant, reflecting the damaging effect of incompatible segmentation.

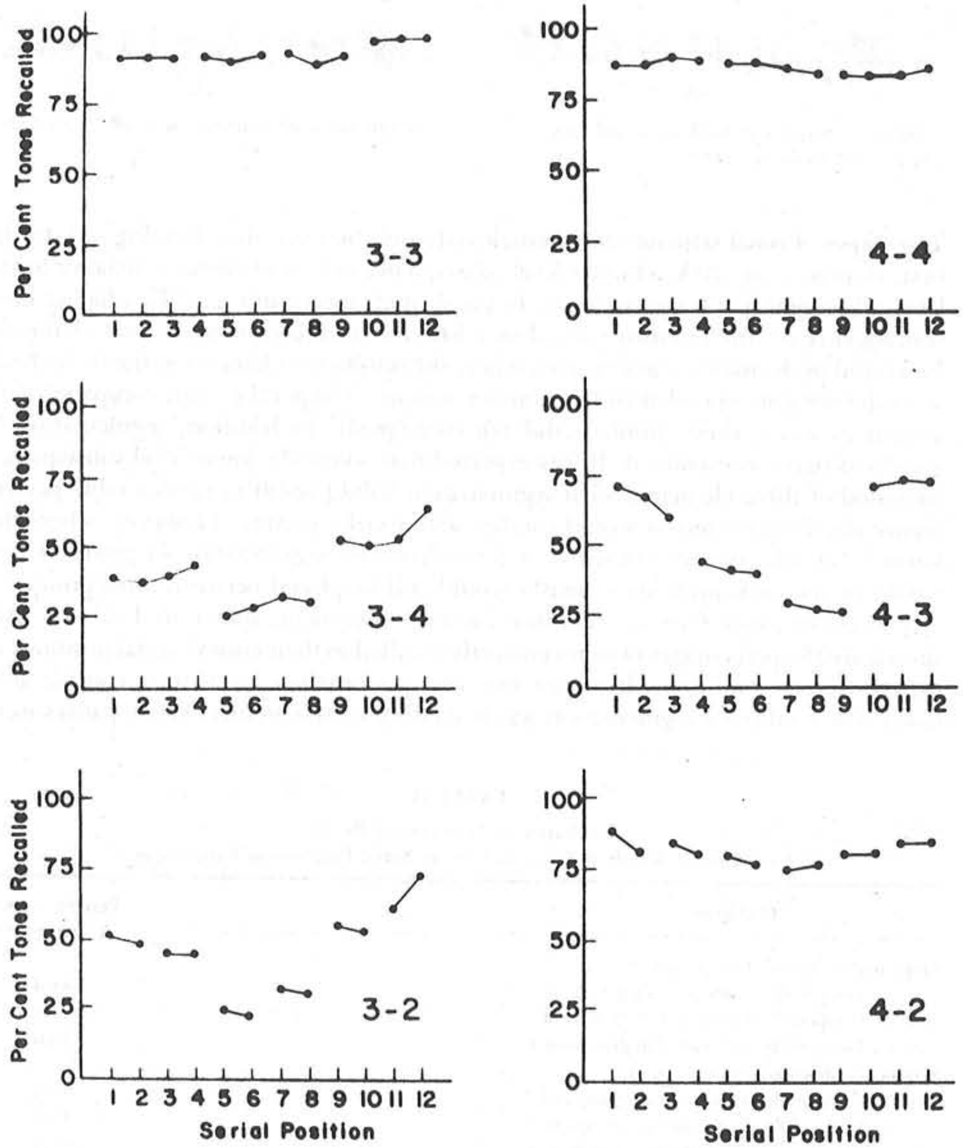

Fig. 21. Serial position curves for the different conditions of the second experiment to study utilization of structure in recall (from Deutsch, 1981).

Figure 21 displays the percentages of tones correctly recalled at each serial position in the different conditions of the experiment. It can be seen that again, discontinuities appear at temporal group boundaries, reflecting the formation of subjective chunks on the basis of temporal proximity.

These two experiments lead to several conclusions. First, they demonstrate that listeners perceive hierarchical structures that are present in tonal sequences and can utilize these structures in recall. For the structured sequences employed in this study, the listener need only retain two chunks of three or four items each; but for the unstructured sequences no such parsimonious encoding was possible. The unstructured sequences therefore imposed a much heavier memory load, with resultant performance decrements. Second, the experiments demonstrate that temporal segmentation has a profound effect on perceived structure, as has been shown by others with different materials. On our present line of reasoning, when segmentation is in conflict with structure, there results a less parsimonious representation, which in turn leads to decrements in recall.

For example, Sequence 2 in Fig. 20 would be encoded in the absence of temporal segmentation as

$$\begin{aligned} A &= (3n,*)G \\ B &= (2n,*)G \\ S &= A[pr]B,G_4 \end{aligned}$$

However, with temporal segmentation in groups of four, the same sequence would be encoded as

$$\begin{aligned} A &= (2n,*)G_{tr} \\ B &= (2n,p,*)G \\ C &= (n,p,n,*)G \\ D &= (p,2n,*)G \\ S &= A[pr](B,C,D),G_4 \end{aligned}$$

Thus, four structures would need to be encoded and retained, together with their alphabets and rules of combination. A much heavier memory load would therefore be imposed.

VI. CONCLUSION

In the foregoing pages, we have considered the rules whereby abstractions based on pitch are formed and also how pitch information is retained at the different levels of abstraction. Where appropriate, we have considered underlying neurophysiological mechanisms, and we have also attempted to draw on insights provided by music theorists. The system that we are dealing with is clearly very complex, but an understanding of its operation is slowly developing.[6]

[6]This work was supported by United States Public Health Service Grant MH-21001.

REFERENCES

Attneave, F., & Olson, R. K. Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, *84*, 147–165.

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

Babbitt, M. The structure and function of musical theory. *College Music Symposium 5*, 1965, 10–21.

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

Bachem, A. Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, 1954, *26*, 751–753.

Baird, J. W. Memory for absolute pitch; Studies in psychology. In *Titchener commemorative volume*. Worcester, 1917. Pp. 69.

Benjamin, W. E. Review of Forte's *The structure of atonal music. Perspectives of New Music*, 1974, *2*, 170–211.

Blackwell, H. R., & Schlosberg, H. Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology*, 1943, *33*, 407–419.

Boomslitter, P., & Creel, W. The long pattern hypothesis in harmony and hearing. *Journal of Music Theory*, 1961, *5*, No. 2, 2–30.

Bower, G. H., & Springston, F. Pauses as recoding points in letter series. *Journal of Experimental Psychology*, 1970, *83*, 421–430.

Bower, G., & Winzenz, D. Group structure, coding and memory for digit series. *Journal of Experimental Psychology Monographs*, 1969, *80*, No. 2, Part 2, 1–17.

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

Browne, R. Review of *The structure of atonal music* by A. Forte. *Journal of Music Theory*, 1974, *18*, 390–415.

Carterette, E. C., Friedman, M. P., & Lovell, J. D. Mach bands in hearing. *Journal of the Acoustical Society of America*, 1969, *45*, 986–998.

Carterette, E. C., Friedman, M. P., & Lovell, J. D. Mach bands in auditory perception. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Sijthoff: Leiden, 1970.

Chrisman, R. Identification and correlation of pitch-sets. *Journal of Music Theory*, 1971, *15*, 58–83.

Collins, A. M., & Quillian, M. R. How to make a language user. In E. Tulving and W. Donaldson (Eds.). *Organization of memory*. New York: Academic Press, 1972.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 671–684.

Cuddy, L. L., & Cohen, A. J. Recognition of transposed melodic sequences. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 255–270.

Deese, V., & Grindley, G. C. The transposition of visual patterns. *British Journal of Psychology*, 1947, *37*, 152–163.

Deutsch, D. Music recognition. *Psychological Review*, 1969, *76*, 300–307.

Deutsch, D. The deterioration of pitch information in memory. Unpublished doctoral dissertation, University of California at San Diego, 1970. (a)

Deutsch, D. Tones and Numbers: Specificity of interference in short-term memory. *Science*, 1970, *168*, 1604–1605. (b)

Deutsch, D. Octave generalization and tune recognition. *Perception & Psychophysics*, 1972, *11*, 411–412. (a)

Deutsch, D. Mapping of interactions in the pitch memory store. *Science*, 1972, *175*, 1020–1022. (b)

Deutsch, D. Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology*, 1972, *93*, 156–162. (c)

Deutsch, D. Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics*, 1973, *13*, 271–175. (a)

Deutsch, D. Interference in memory between tones adjacent in the musical scale. *Journal of Experimental Psychology*, 1973, *100*, 228–231. (b)

Deutsch, D. Generality of interference by tonal stimuli in recognition memory for pitch. *Quarterly Journal of Experimental Psychology*, 1974, *26*, 229–234.

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

Deutsch, D. The organization of short-term memory for a single acoustic attribute. In D. Deutsch & J. A. Deutsch (Eds.), *Short term memory*. New York: Academic Press, 1975. Pp. 107–151. (b)

Deutsch, D. Facilitation by repetition in recognition memory for tonal pitch. *Memory & Cognition*, 1975, *3*, 263–266. (c)

Deutsch, D. Auditory memory. *Canadian Journal of Psychology*, 1975, *29*, 87–105. (d)

Deutsch, D. Octave equivalence and the processing of melodic sequences. *Journal of the Acoustical Society of America*, 1976, *60*, S94.

Deutsch, D. Memory and attention in music. In M. Critchley & R. A. Henson (Eds.). *Music and the brain*, London: Heinemann, 1977. Pp. 95–130.

Deutsch, D. Octave generalization and melody identification. *Perception & Psychophysics*, 1978, *23*, 91–92. (a)

Deutsch, D. Interactive effects in memory for harmonic intervals. *Perception & Psychophysics*, 1978, *24*, 7–10. (b)

Deutsch, D. Delayed pitch comparisons and the principle of proximity. *Perception & Psychophysics*, 1978, *23*, 227–230. (c)

Deutsch, D. The psychology of music. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*. (Volume X) New York: Academic Press, 1978. Pp. 191–218. (d)

Deutsch, D. Octave generalization and the consolidation of melodic information. *Canadian Journal of Psychology*, 1979, *33*, 201–204.

Deutsch, D. Recall of hierarchically structured melodic sequences. *Journal of the Acoustical Society of America*, 1980, *67*, 57.

Deutsch, D. The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, 1981, *28*, 381–389.

Deutsch, D. Facilitation and misrecognition effects in pitch memory as a function of serial position and temporal proximity. In preparation.

Deutsch, D. Effects of melodic context on pitch recognition judgment. *Perception and Psychophysics*, in press.

Deutsch, D., & Feroe, J. Disinhibition in pitch memory. *Perception & Psychophysics*, 1975, *17*, 320–324.

Deutsch, D., and Feroe, J. The internal representation of pitch sequences in tonal music. *Psychological Review*, 1981, *88* 503–522.

Deutsch, D., & Roll, P. L. Error patterns in delayed pitch comparison as a function of relational context. *Journal of Experimental Psychology*, 1974, *103*, 1027–1034.

Deutsch, J. A., & Deutsch, D. *Physiological psychology*. (2nd Ed.) Homewood: Dorsey, 1973.

Dewar, K. M. Context Effects in Recognition Memory for Tones. Unpublished doctoral dissertation. Queens University, Kingston, Ontario, 1974.

Dewar, K. M., Cuddy, C. L., & Mewhort, D. J. K. Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, *3*, 60–67.

Divenyi, P. L., & Hirsh, I. J. Some figural properties of auditory patterns. *Journal of the Acoustical Society of America*, 1978, *64*, 1369–1385.

Dowling, W. J. Rhythmic Fission and the Perceptual Organization of Tone Sequences. Unpublished doctoral dissertation, Harvard University, 1967.

Dowling, W. J. Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics*, 1973, *4*, 37–40.

Dowling, W. J. Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 1978, *85*, 342–354.

Dowling, W. J., & Fujitani, D. S. Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 1971, *49*, 524–531.

Dowling, W. J., & Hollombe, A. W. The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception & Psychophysics*, 1977, *21*, 60–64.

Drobisch, M. W. *Uber die mathematische Bestimmung der musikalischen*. Intervalle, 1846.

Drobisch, M. W. *Abhandlungen der Sachsischen Wissenchaften math-phys.* C1.B.II. 1855, 35.

Erickson, R. Paper presented to Department of Music, University of California, San Diego, California, 1978.

Ernst, G. W., & Newell, A. *GPS: A Case Study in Generality and Problem Solving.* New York: Academic Press, 1969.

Evans, E. F. Neural processes for the detection of acoustic patterns and for sound localization. In F. O. Schmitt and F. T. Worden (Eds.), *The neurosciences, third study program.* Cambridge, Massachusetts: MIT Press, 1974. Pp. 131-147.

Forte, A. *The structure of atonal music,* New Haven, Connecticut: Yale University Press, 1973.

Francès, R. *La perception de de musique.* Paris, Vrin, 1958.

Fucks, W. Mathematical analysis of the formal structure of music. *Institute of Radio Engineers Transactions on Information Theory,* 1962, *8,* 225-228.

Gagné, R. M. The acquisition of knowledge. *Psychological Review,* 1962, *67,* 355-365.

Garner, W. R. *The processing of information and structure.* Hillsdale, New Jersey: Erlbaum, 1974.

Greeno, J. G., & Simon, H. A. Processes for sequence production. *Psychological Review,* 1974, *81,* 187-196.

Handel, S. Temporal segmentation of repeating auditory patterns. *Journal of Experimental Psychology,* 1973, *101,* 46-54.

Hanson, A. R., & Riseman, E. M. (Eds.) *Computer vision systems.* New York: Academic Press, 1978.

Harris, J. D. The decline of pitch discrimination with time. *Journal of Experimental Psychology,* 1952, *43,* 96-99.

Helmholtz, H. Von. *On the sensations of tone as a physiological basis for the theory of music.* (2nd English ed.) New York: Dover, 1954. (Originally published 1859).

Helmholtz, H. Von. The Origin of the Correct Interpretations of our Sensory impressions. *Zeitchrift für Psychologie und Physiologie der Sinnesorgane* 1894, 7, 81-96.

Houtgast, T. Psychophysical evidence for lateral inhibition in hearing. *Journal of the Acoustical Society of America,* 1972, *51,* 1885-1894.

Howe, H. S. Some combinatorial properties of pitch structures. *Journal of Music Theory,* 1965, *4,* 45-61.

Howe, H. S. Review of Forte's *The structure of atonal music. Collectanea,* 1974, 118-124.

Hubel, D. H., & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology,* 1962, *160,* 106-154.

Humphreys, L. F. Generalization as a function of method of reinforcement. *Journal of Experimental Psychology,* 1939, *25,* 361-372.

Idson, W. L., & Massaro, D. W. A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics,* 1978, *24,* 551-565.

Jeffries, T. B. Relationship of interval frequency count to ratings of melodic intervals. *Journal of Experimental Psychology,* 1974, *102,* 903-905.

Jones, M. R. Cognitive representations of serial patterns. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, New Jersey: Erlbaum, 1974.

Jones, M. R. Auditory patterns: Studies in the perception of structure. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception.* (Vol. VIII, Perceptual coding) New York: Academic Press, 1978.

Julesz, B., & Hirsh, I. J. Visual and auditory perception—An essay of comparison. In E. E. David and P. B. Denes (Eds.), *Human communication: A unified view.* New York: McGraw-Hill, 1972.

Kallman, H. J., & Massaro, D. W. Tone chroma is functional in melody recognition. *Perception & Psychophysics,* 1979, *26,* 32-36.

Klinke, R., Boerger, G., & Gruber, J. Alteration of afferent, tone-evoked activity of neurons of the cochlear nucleus following acoustic stimulation of the contralateral ear. *Journal of the Acoustical Society of America,* 1969, *45,* 788-789.

Klinke, R., Boerger, G., & Gruber, J. 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, 1970.

Koester, T. The time error in pitch and loudness discrimination as a function of time interval and stimulus level. *Archives of Psychology,* 1945, *297,* whole issue.

Koffka, K. *Principles of Gestalt psychology.* New York, Harcourt, 1935.

Krumhansl, C. L. The psychological representation of musical pitch in a tonal context. *Cognitive Psychology,* 1979, *11,* 346-374.

Leeuwenberg, E. L. A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, 1971, *84*, 307–349.

Lerdahl, F., & Jockendorff, R. Toward a formal theory of music. *Journal of Music Theory* 1977, *21*, No. 2, 111–172.

Lewin, D. The intervallic content of a collection of notes. *Journal of Music Theory*, 1960, *4*, 98–101.

Lewin, D. A theory of segmental association in twelve-tone music. *Perspectives of New Music*, 1962, *1*, 89–116.

Longuet-Higgins, H. C. Letter to a musical friend. *Music Review*, 1962, *23*, 244–248. (a)

Longuet-Higgins, H. C. Second letter to a musical friend. *Music Review*, 1962, *23*, 271–280. (b)

Longuet-Higgins, H. C. The perception of music. *Interdisciplinary Science Reviews*, 1978, *3*, 148–156.

Mach, E. *The analysis of sensations and the relation of the physical to the psychical*. (C. M. Williams, translator; W. Waterlow, review and supplement) New York: Dover, 1959. (Originally published in German, 1906)

McLean, R. S., and Gregg, L. W. Effects of induced chunking on temporal aspects of serial retention. *Journal of Experimental Psychology*, 1967, *74*, 455–459.

Merriam, A. P. *The anthropology of music*. Evanston, Illinois: Northwestern University Press, 1964.

Meyer, L. B. *Emotion and meaning in music*. Chicago, Illinois: University of Chicago Press, 1956.

Meyer, L. B. *Explaining music: Essays and explorations*. Berkeley, California: University of California Press, 1973.

Meyer, M. On the attributes of the sensations. *Psychological Review*, 1904, *11*, 83–103.

Meyer, M. Review of G. Révész, "Zur Grundleguncy der Tonpsychologie." *Psychological Bulletin*, 1914, *11*, 349–352.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, *63*, 81–97.

Miller, G. A., & Chomsky, N. Finitary models of language users. *Handbook of Mathematical Psychology*, 1963, *2*, 419–493.

Miller, G. A., Galanter, E. H., & Pribram, K. H. *Plans and the structure of behavior*. New York: Holt, 1960.

Mueller, G. E., & Schumann, F. Experimentelle Beitrage zur Untersuchung des Gedächtnisses. *Zeitschrift fur Psychologie und Physiologie der Sinnesorgane*, 1894, *6*, 81–190, 257–339.

Narmour, E. *Beyond Schenkerism*. Chicago, Illinois: University of Chicago Press, 1977.

Navon, D. Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 1977, *9*, 353–383.

Nettl, B. *Music in primitive culture*. Cambridge, Massachusetts: Harvard University Press, 1956.

Olson, R. K., & Hanson, V. Interference effects in tone memory. *Memory & Cognition*, 1977, *5*, 32–40.

Ortmann, O. On the melodic relativity of tones. *Psychological Monographs*, 1926, *35*, Whole No. 162.

Palmer, S. E. The effects of contextual scenes on the identification of objects. *Memory and Cognition*, 1975, *3*, 519–526.

Perle, G. *Serial composition and atonality*. (3rd ed.) Berkeley, California: University of California Press, 1972.

Perle, G. *Twelve-tone tonality*. Berkeley, California: University of California Press, 1977.

Piston, W. *Harmony*. (2nd ed.) London: Norton, 1948.

Pitts, W., & McCulloch, W. S. How we know universals. The perception of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 1947, *9*, 127–147.

Plomp, R., Wagenaar, W. A., & Mimpen, A. M. Musical interval recognition with simultaneous tones, *Acustica*, 1973, *29*, 101–109.

Ratliff, F. *Mach bands: Quantitative studies of neural networks in the retina*. San Francisco, California: Holden Day, 1965.

Restle, F. Theory of serial pattern learning: Structural trees. *Psychological Review*, 1970, *77*, 481–495.

Restle, F. Serial patterns: The role of phrasing. *Journal of Experimental Psychology*, 1972, *92*, 385–390.

Restle, F., & Brown, E. Organization of serial pattern learning. In G. H. Bower (Ed.), *The psychology of learning and motivation*. (Volume 4) New York: Academic Press, 1970. Pp. 249–331.

Révész, G. *Zur grundleguncy der tonpsychologie*. Leipzig: Feit, 1913.

Ruckmick, C. A. A new classification of tonal qualities. *Psychological Review*, 1929, *36*, 172–180.

Salzer, F. *Structural hearing*. New York: Dover, 1962.

Scandura, J. M. Role of rules in behavior: Toward an operational definition of what (rule) is learned. *Psychological Review*, 1970, 77, 516–533.

Schenker, H. *Neue musikalische theorien und phantasien: Der Freie Satz*. Vienna: Universal Edition, 1956.

Schenker, H. In O. Jonas (Ed. & Annotating), *Harmony*. (E. M. Borgese, Translator). Cambridge, Massachusetts: MIT Press, 1973.

Schoenberg, A. *Style and idea*. London: Williams & Norgate, 1951.

Schoenberg, A. In G. Strong (Ed.), *Fundamentals of musical composition*. New York: St. Martin's Press, 1967.

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

Simon, H. A. Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 1972, *79*, 369–382.

Simon, H. A., & Kotovsky, K. Human acquisition of concepts for sequential patterns. *Psychological Review*, 1963, *70*, 534–546.

Simon, H. A., & Sumner, R. K. Pattern in music. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: Wiley, 1968.

Suga, N., O'Neill, W. E., & Manabe, T. Harmonic-sensitive neurons in the auditory cortex of the mustache bat. *Science*, 1979, *203*, 270–274.

Suga, N., & P. H-S. Jen. Disproportionate tonotopic representation for processing CF-FM sonar signals in the mustache bat auditory cortex. *Science*, 1976, *194*, 542–544.

Teitelbaum, R. Intervallic relations in atonal music. *Journal of Music Theory*, 1965, *9*, 72–127.

Van Noorden, L. P. A. S. Temporal coherence in the perception of tone sequences. Unpublished doctoral thesis. Technische Hogeschool, Eindhoven, Holland, 1975.

Vitz, P. C., & Todd, T. C. A model of learning for simple repeating binary patterns. *Journal of Experimental Psychology*, 1967, *75*, 108–117.

Vitz, P. C., & Todd, T. C. A coded element model of the perceptual processing of sequential stimuli. *Psychological Review*, 1969, *76*, 433–449.

Von Ehrenfels, C. *Uber Gestaltqualitäten Vierteljahrschrift fur Wissenschaftliche Philosophie*, 1890, *14*, 249–292.

Warren, R. M., & Warren, R. P. *Helmholtz on perception: Its physiology and development*. New York: Wiley, 1968.

Werner, H. Uber Mikromelodik und Mikroharmonik. *Zeitschrift fur Psychologie*, 1925, *98*, 74–89.

Wever, E. G., & Bray, C. W. The nature of the acoustic response; the relation between sound frequency and frequency of impulses in the auditory nerve. *Journal of Experimental Psychology*, 1930, *13*, 373–387.

White, B. Recognition of distorted melodies. *American Journal of Psychology*, 1960, *73*, 100–107.

Wickelgren, W. A. Consolidation and retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, 1966, *72*, 250–259.

Wickelgren, W. A. Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology*, 1969, *6*, 13–61.

Winston, P. H. Learning to identify toy block structures. In R. L. Solo (Ed.), *Contemporary issues in cognitive psychology: the Loyola Symposium*. Washington, D. C.: Winston, 1973.

Yeston, M. (Ed.) *Readings in Schenker analysis and other approaches*. New Haven, Connecticut: Yale University Press, 1977.

Yngve, V. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 1960, *104*, 444–466.