Chapter 9

THE PSYCHOLOGY OF MUSIC*

DIANA DEUTSCH

I. INTRODUCTION

The perceptual psychologist will find in music a rich and rewarding area for investigation. Here he can explore many problems, such as shape recognition, attention, memory, and even abstract cognitive activity, generally without the complication of verbal labeling by the subject. Furthermore, he can employ complex stimuli derived from music such as sequences of tones, or complex sounds of varying timbre, to broaden our understanding of the human auditory system beyond that which can be derived from classical psychophysical stimuli.

---

* This work was supported by United States Public Health Service Grant MH-21001-03.

HANDBOOK OF PERCEPTION, VOL. X

ISBN 0-12-161910-9

E. C. Carterette and M. P. Friedman (Eds.)

In the past, the study of musical information processing has been hampered by technical difficulties involved in generating complex auditory stimuli with precisely controlled parameters. However, the technological advances of recent years have completely changed this picture; as a result, experimental interest in this field has developed rapidly. Given the intrinsic interest of the subject, and the fact that it is still largely unexplored, we may expect to see considerable advances in our understanding of musical psychology over the next decade.

Due to lack of space, this chapter omits certain topics that should be included in a comprehensive account of the psychology of music. Questions involving the origins and functions of music are not discussed, since these are not fundamentally problems in perception. For accounts of basic psychoacoustical phenomena and the physiology of the auditory pathway, the reader is referred to Volume IV of this *Handbook*. A comprehensive discussion of tuning systems, and of consonance and dissonance, is to be found in Ward (1970).

II. RECOGNITION OF TONAL SHAPE

The abstraction of tonal information occurs in several stages. The initial stages of abstraction involve the detection of both specific features (such as intervals, chords, and tone chroma) and global features (such as contour). Such features are then combined to form tonal sequences, which are themselves abstracted and combined in systematic ways. These stages of abstraction will be discussed in what follows.

A. Tonal Features

1. TONE CHROMA

The term *tone chroma* refers to the position of a tone within the octave. Tones that are separated by octaves are given the same name in traditional Western music, and are treated as perceptually similar. Indeed, people with absolute pitch sometimes place a note in the wrong octave even though they name it correctly (Baird, 1917; Bachem, 1954). Octave duplications in scales are also found in other cultures (Nettl, 1956a). Further evidence for the perceptual equivalence of tones separated by octaves comes from conditioning studies. Generalization of response to tones placed an octave apart occurs both in man (Humphreys, 1939) and in animals (Blackwell & Schlosberg, 1943). Such octave generalization has been used to produce a compelling illusion. Shepard (1964) generated a set of complex tones, each of which consisted of several tones separated by octaves. When these tones were presented in ascending semitonal steps at a rate of about 1 sec^{-1}, listeners perceived a set of tones that constantly increased in pitch and never returned to the beginning. These tones therefore appeared to be endlessly climbing up an abstracted octave.

The physical octave represents a frequency ratio of 2:1. However, as first noted by Stumpf and Meyer (1898), the subjective octave is slightly larger than this. Ward (1954) made a systematic investigation of this phenomenon. He presented subjects with two pure tones in repetitive succession, and required them to adjust the frequency of one to be exactly an octave above the other. It was found that subjects made adjustments that were reliably larger than a 2:1 ratio and, further, that the amount of deviation from the physical octave increased in the higher frequency ranges. Sundberg and Lindquist (1973) obtained similar findings with complex tones. Furthermore, Burns (1974a) repeated Ward's experiment with professional Indian musicians as subjects, and obtained remarkably similar results; so it would appear that the phenomenon of octave stretch is not dependent on culture. This conclusion is reinforced by the finding that tuning practices in several non-Western cultures agree very well with these laboratory measurements (Dowling, 1973a). Terhardt (1971) proposed that the phenomenon of octave stretch is acquired early in life as a result of exposure to complex sounds. In such sounds, the pitches of neighboring partials move slightly away from each other, as a result of a mutual masking effect, and Terhardt assumed that we generalize from this to harmonically related tones that are presented in succession. On the other hand, the phenomenon may simply reflect innate properties of the auditory system—an explanation favored by Dowling (1973a).

Because of the perceptual similarity of tones separated by octaves, it makes most sense to regard pitch as a bidimensional attribute: one dimension being a monotonic function of frequency, or tone height (at least in the case of pure tones), and the other defining the position of a tone within the octave, or tone chroma. Various representations of pitch have been proposed that accommodate these two dimensions. Drobisch (1846) suggested that pitch be represented as a line spiraling around a helix, with tones separated by an octave lying most proximal within each turn of the helix. Ruckmick (1929) made an essentially similar proposal involving a bell-shaped spiral instead. However, such suggestions do not make plausible neurophysiological models. It would seem more in line with our knowledge of brain function to suggest that neural units underlying tones that are separated by octaves converge onto the same higher-order units (Deutsch, 1969). Indeed, units with two or more "best" frequencies that are octave multiples of each other have been found in the auditory system (Evans, 1974), and such units are very likely to receive multiple projections from units with single best frequencies.

2. INTERVALS AND CHORDS

The simultaneous or successive presentation of two tones results in the perception of a musical interval; and intervals whose components are separated by the same frequency ratio are perceived as being the same size. That is, tone pairs F_1 and F_2 and pairs F_3 and F_4 are perceived as equivalent when $\log F_1 - \log F_2 = \log F_3 - \log F_4$. This principle forms an important basis for the traditional musical scale. The smallest unit of this scale is the semitone, which represents a frequency ratio of approximately 18 : 17; tone pairs that are separated by the same number of semitones are given the same name. Furthermore, although music in other cultures may involve intervals of different sizes, the perceptual equivalence of tone pairs that stand in the same frequency ratio appears to hold cross-culturally (Nettl, 1956a). Attneave and Olson (1971) made a careful study of transposition in a laboratory situation. They presented subjects with a simple melodic pattern, and then required them to reproduce this pattern in a different frequency range. It was found that when the subjects were able to draw on information that was well embedded in long-term memory, they clearly transposed on a log-frequency medium.

Chords of three or more tones are also classified on the basis of the frequency ratios of their components; those standing in the same frequency relationship are given the same name. However, this perceptual equivalence is not based simply on an identity in the set of the component intervals. For instance, a major triad and a minor triad sound quite different, yet both are composed of a major third, a minor third, and a fifth. However, these occur in a different order. In the major triad the major third lies below the minor third; in the minor triad the minor third lies below the major third. Deutsch (1969) proposed a model for interval and chord abstraction that accommodates this feature. This model assumes that such abstractions take place in two stages. At the first stage, units that respond to tones of specific pitch converge in groups of two or three onto the same second-order unit. Such second-order units then converge onto third-order units in such a way that all those activated by tonal combinations standing in the same pitch relationship are linked together. These third-order units therefore respond to intervals and chords irrespective of the pitches of their components.

Another shape-recognition operation is also performed in the case of simultaneous intervals and chords. When the components of such combinations are placed in different octaves, the combinations retain a perceptual similarity. This phenomenon is known in music as inversion. Thus, a simultaneous interval of n semitones is perceptually similar to a simultaneous interval of $12 - n$ semitones. Plomp, Wagenaar, and Mimpen (1973) provided evidence for this perceptual similarity in a laboratory situation. Subjects were required to identify intervals formed by simultaneous tone pairs, and it was found that confusions occurred between intervals that were inversions of each other. In another experiment by Deutsch and Roll (1974), subjects made pitch-comparison judgments when the tones to be compared were accompanied by tones of lower pitch. Misrecognition errors were quite pronounced when the standard and comparison tones differed, but the standard and comparison combinations formed identical intervals. Misrecognition errors were also found to increase significantly when the interval formed by the standard combination was an inversion of the interval formed by the comparison combination. It was concluded that the subjects were basing their false recognition judgments on the perceptual equivalence of the identical or inverted intervals.

3. CATEGORICAL PERCEPTION OF MUSICAL INTERVALS

As pointed out by Francès (1972), the perceptual elements of music in actual practice are not tones, but rather notes, which are generally very imprecisely emitted. In singing, and also in playing many musical instruments, notes are produced with a vibrato involving substantial frequency modulation. In one experiment, Francès (1972) found that the average range of vibrato for three professional singers was over a semitone. Furthermore, the average pitch of a note generated by voice or by natural instruments often changes during its presentation, and in passing from one note to another, glides through intervening frequencies are common. However, when we listen to music, we generally ignore these irregularities and instead form abstractions out of the tonal elements and the relationships between them. In one study, Burns and Ward (1973) required musicians to make identification and discrimination judgments involving musical intervals and found the same type of categorical perception as has been found for speech stimuli (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Similar results were obtained by Siegel and Sopo (1975).

There are various reasons why we should form relatively broad categories in perceiving tonal sequences. First, although the difference limen for pitch, as measured by comparing two pitches in isolation, is well below a semitone, this is drastically raised when the tones to be compared are separated by a sequence of intervening tones (Deutsch, 1970a). The formation of relatively broad categories may therefore be necessary in order for the listener to recognize repeated elements. Second, gross inaccuracies occur in emitting notes by voice and by many natural instruments, and such inaccuracies must be ignored if the listener is to identify repetitions. If this line of reasoning is correct, we should expect to find a certain minimum of interval size to occur cross-culturally. This is in fact the case. Nettl (1956a) concluded upon reviewing scales in primitive cultures that intervals smaller than a semitone are very rare. When such intervals do occur, this is almost always under specialized conditions, such as in ornamentation. For instance, the scale in Hindustani music is theoretically divided into 22 intervals (*shrutis*) to the octave. However, in actual practice, two tones separated by one *shruti* occur only under special circumstances (Bake, 1957). It is particularly interesting in this regard to find that professional Indian musicians, trained in Hindustani music, were unable to identify musical intervals with any more precision than Western musicians (Burns, 1974b), and indeed did not identify *shruti*s with any consistency in the laboratory.

Despite such broad categorization, there do appear to be preferred tunings in at least some musical situations. For a detailed discussion of tuning preferences, the reader is referred to Boomsliter and Creel (1963) and Ward (1970).

4. CONTOUR

The tonal information in a piece of music can also be described in global fashion; for instance, in terms of its range, the proportion and average size of ascending compared with descending intervals, the sequence of directions of pitch change, and so on. There is evidence that such global cues are used in recognizing music. Early psychologists, such as Werner (1925), showed that people could recognize familiar melodies when these were transformed onto very small scales, so that the actual intervals were grossly distorted in size. More recently, White (1960) required subjects to identify well-known melodies that were distorted in various ways. Some recognition was still obtained when all the intervals were set to one semitone, showing that we can recognize a melody on the basis of the sequence of directions of pitch change alone. However, performance was substantially improved when the relative sizes of the intervals were left unchanged, even though their absolute sizes were altered.

The importance of contour in short-term melody recognition is shown by the fact that in actual music, phrases are often repeated under transposition with distortions in the interval sizes but with contour intact (see Fig. 1). In an experiment by Dowling and Fujitani (1971), subjects made short-term recognition judgments involving transposed melodies. The subjects were very proficient at recognizing melodies that were either transposed, but identical, or where the contour alone was preserved. However, they appeared incapable of discriminating between exact transpositions and those that were merely preserved contour. Yet, when long-term memory for familiar tunes was tested instead, recognition on the basis of contour alone was quite unimpressive, so it appears that consolidation of memory for melodies acts on specific interval information, rather than on contour. This may be related to the finding by Attneave and Olson (1971) that some subjects showed considerable variability in transposing tonal sequences in a short-term memory situation, yet clearly transposed on a log-frequency continuum when they were given a highly overlearned sequence as the standard.

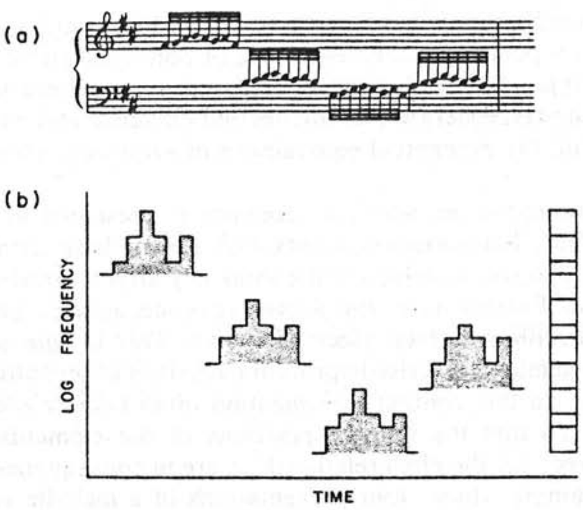

FIG. 1. Transposition along the scalar alphabet. The same melodic configuration is presented four times in succession, at different positions along the scale. In consequence, there is a variation in the set of intervals involved. The sequence in musical notation is shown in (a). The sequence plotted as log frequency versus time is shown in (b). The ladder at the right shows the scale. [(a) From J. S. Bach, *The well-tempered clavier, Book I, Fugue V.* (b) From D. Deutsch, *Memory and attention in music*. In M. Critchley & R. A. Henson (Eds.). *Music and the Brain.* London: Heinemann Medical Books, Ltd., 1977. Reprinted by permission.]

B. Scales

In all cultures, tonal sequences are formed out of small, fairly well-defined sets of pitch relationships, which constitute their scale. For any given scale, one can also determine a mode (i.e., a set of a priori probabilities of occurrence for notes standing in different positions along the scale) and a set of *transitional* probabilities between the different notes. This is true both for linear successions of tones and also for harmonic sequences. In traditional Western harmony there are very strong *transitional* probabilities governing the root progressions of chords (Piston, 1948). As pointed out by Meyer (1956), the use of scales with strong a priori and transitional probabilities contributes substantially to the ease of musical processing, since it enables us to draw on highly overlearned information in listening to unfamiliar sequences.

C. Recognition of Tonal Sequences under Various Transformations

The question of how well we recognize a tonal sequence when it has undergone a given transformation is analogous to asking how well a visual shape is recognized when it has been transformed in some way, such as changing its size or position in the visual field or converting it into its mirror image. Various types of transformation are used in actual music. Some of these are spontaneous, others are an intellectual exercise, and musicologists disagree widely on the perceptual equivalence of sequences transformed in various ways.

The case of transposition, where a sequence is presented in a different pitch range, is clear. Transposition occurs so readily in long-term recall that it is rare to be able to sing a melody in the same key after a certain amount of time has elapsed. Furthermore, the same sequence is often presented in transposed form within a given piece of music. This is true not only in Western cultivated music, but also in primitive music of other cultures (Nettl, 1956b). However, in this context transposition often takes place along the scalar alphabet, so that the relative positions of the elements along the scale are preserved, but the pitch relationships are in consequence distorted. Figure 1, for example, shows four presentations of a melodic sequence at different points along the scale, involving three different sets of pitch relationships.

Other transformations performed on linear tonal sequences are known as *inversion* and *retrogression*. In inversion (not to be confused with the harmonic use of the term), the directions of the melodic intervals are systematically reversed. That is, all ascending intervals become descending intervals, and vice versa. In traditional music, inversion generally leads to a change in the set of intervals, in order to preserve the set of elements along the scale. In retrogression, the sequence is played backward. This operation is used only rarely in traditional music (Piston, 1949; Tovey, 1957), and then generally with a change in the set of intervals.

Dowling (1972) has raised the possibility that sequences transformed by the operations of retrogression or inversion are recognized on the basis of their contour, rather than by abstracting their actual melodic intervals. He presented listeners first with a standard five-tone melody and then a comparison melody. The second melody was either unrelated to the first, an exact transposition of the first, or was transformed by the operation of inversion, retrogression, or retrograde inversion (a combination of the two). In a further set of conditions, the comparison sequence under these various transformations was further distorted so that its contour was preserved but the exact intervals were destroyed. Dowling found that although subjects could recognize a transformed sequence with better than chance accuracy, there was no evidence that they distinguished between transformations that preserved the intervals and those that preserved contour alone.

In another study, White (1960) investigated long-term recognition of melodies played in retrogression. He found that although subjects showed some recognition of such melodies, performance was at about the same level as when the melody was played in a monotone with the rhythmic information alone remaining. Even more interestingly, recognition was better for sequences where the intervals within the melody were randomly permuted than for straight temporal reversal. It would appear, therefore, that the subjects were identifying the reversed sequences on the basis of their component intervals alone, rather than on the ordering of these intervals.

A further transformation to consider involves an extension of the principle of octave equivalence to tonal sequences. One might assume that, because of the perceptual similarity of tones separated by octaves, a sequence would also remain perceptually equivalent if its components were placed in different octaves. This question has been experimentally investigated in a long-term memory situation. Deutsch (1972a) presented subjects with the first half of the tune "Yankee Doodle" under various conditions. This tune was universally recognized when presented in any one of three octaves. However, when the notes of the tune were chosen randomly from these same three octaves with the restriction that no two successive notes occurred as in the untransformed version, recognition was no better than when the sequence was played as a series of clicks with the pitch information removed but the rhythm retained. Dowling and Hollombe (1977) obtained similar results using several different tunes. They also found that leaving the melodic contour intact improved performance for subjects informed of the preservation of contour, though performance was still substantially worse than for the untransformed versions. This is in accordance with the evidence described above that contour alone can form a basis for melody recognition.

Deutsch (1972a) also found that when the listeners were informed of the indentity of the tune, they had no trouble in following the randomized-octaves version and confirming that each note was correctly placed within its octave. One can suppose that this was achieved by the subjects imagining the tune to themselves as they heard the randomized-octaves version and matching each note as it arrived with its octave equivalent. Thus, we would expect that recognition of a note when it is displaced to another octave would depend on the strength of expectation for this note. The operation could therefore be meaningfully performed, even using a new sequence, if the cognitive structure of the sequence was such that the displaced note was highly probable. Meyer (1973) has also argued that melodic processes are generally specific to a given octave, and that octave jumps should only be made under special conditions.

D. Hierarchical Organization

Human serially patterned behavior is characteristically organized in hierarchical fashion. In language, for instance, words are combined to form phrases, which in turn combine to form sentences, and so on. Similarly in music, notes are combined to form motives, which in turn form phrases, and these are themselves hierarchically structured.

Restle and Brown (1970) provide convincing evidence that such hierarchical organization of sequential elements is a general cognitive phenomenon. They presented subjects with a row of lights that turned on and off in repetitive succession. The subjects' task was to predict at each point which light would come on next. It was concluded from the speed with which different types of sequences were learned and also from the patterns of error at different points in a sequence that subjects approach this task by grouping the sequence into a series of chunks (in this case, runs and trills). Restle and Brown then constructed sequences in which such basic chunks were operated on by various rules such as repetition, transposition, and mirror-image reversal, so as to form higher-order chunks. These were in turn transformed by the same types of operations at several levels in such a way as to form structural trees. It was found that sequences structured in this fashion were learned substantially more easily than comparable sequences composed of the same subsequences but lacking this higher-level organization. Furthermore, most errors were found to occur at the highest-order transformations, with the probability of error decreasing as the level of transformation along the tree decreased.

As Restle and Brown (1970) point out, such trees in actual music would have to be substantially more complex than those used in these experiments. For one thing, music involves more than one basic theme. Also, various other operations occur in music, such as interspersing of sequences, and Restle (1973) has demonstrated that subjects can indeed process interspersed sequences in the light-switching task. Furthermore, transposition in a tonal system can involve either shifting the sequence along the scale, or shifting the entire range of the scale, which entails a flexibility of alphabet. In modulating from major to minor there is also a change in alphabet, as is also the case with deviations from the diatonic scale (for instance, to the chromatic scale, or to an alphabet based on an arpeggio). Indeed, different alphabets are often used at different levels of transformation.

An additional complexity lies in the fact that cultivated music involves the presentation of more than one sequence of tones at a time, so that more than one tree structure must be operating simultaneously in processing such music. Furthermore, the harmonic sequences produced by simultaneous tonal combinations are also hierarchically structured, so the choice of notes in each linear row must be such as to satisfy these harmonic requirements as well. Listening to cultivated music must therefore involve parallel processing along several such structural trees (see also Winograd, 1968; Deutsch, 1977).

III. THE FORMATION OF PERCEPTUAL CONFIGURATIONS

When we listen to music, we do not simply process each element as it arrives; rather, we form sequential groupings out of combinations of elements. Once such groupings are formed, there is further a tendency for one to come to the foreground of our attention, while others are relegated to the background. The stability of such figure–ground organization depends on the type of music presented. Thus in some music, such as accompanied songs, one voice tends strongly to be heard as the figure and the other as the ground. In contrast, in contrapuntal music, such as canons and fugues, we attempt as much as possible to attend to all voices, which results in our fluctuating between alternative modes of figure–ground organization.

A. Groupings Based on Pitch Range

The principles governing the formation of perceptual configurations in music may be compared with those in vision (Wertheimer, 1923). One important factor is the law of proximity, especially as applied to frequency range. Thus, when a solo instrument plays a melody and an accompaniment, these are generally in different frequency ranges. Similarly, in contrapuntal music, each voice has its own range. An interesting musical technique that takes advantage of this perceptual tendency is known as *pseudopolyphony*. Here a sequence of notes drawn from two different ranges is played in rapid succession, with the result that the listener perceives two melodic lines in parallel. Examples of such music are shown in Fig. 2.

A number of investigators have studied the perception of rapid sequences of tones as a function of their frequency disparity. Miller and Heise (1950) presented sequences of two alternating tones at a rate of 10 $\sec^{-1}$. They found that when the frequencies of these tones differed by less than 15%, listeners perceived a single string of related tones (i.e., a trill). However, with an increase in the frequency disparity between the alternating tones the sequence was heard as two interrupted and unrelated tones. Heise and Miller (1951) explored this phenomenon in sequences consisting of several tonal frequencies and found that if one of the tones was sufficiently different from the others it was heard in isolation, as though emanating from a different sound source. Van Noorden (1975) studied the effect of different rates of presentation on this phenomenon. He found that when subjects were trying to hear the sequence as coherent, decreasing the presentation rate from 50 to 150 msec for each tone enabled coherence to be heard at substantially larger values of frequency separation. However, when subjects were trying to hear two separate streams, changing the presentation rate had very little effect on performance.

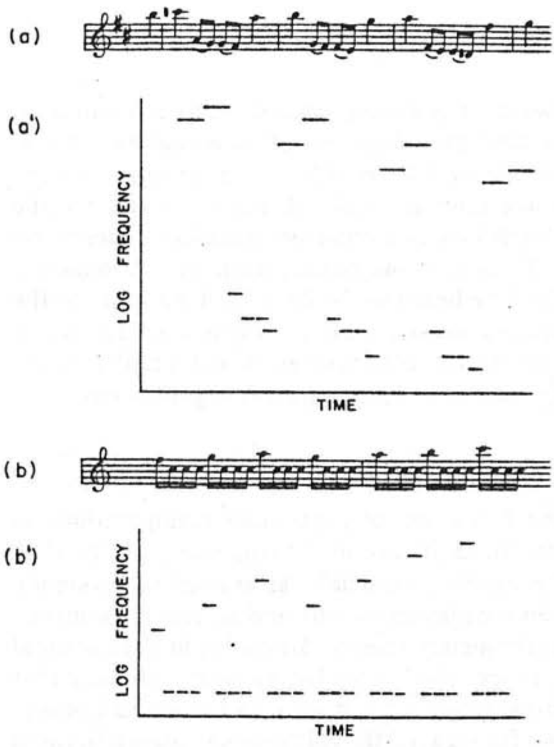

FIG. 2. The grouping of melodic stimuli by frequency proximity. In Sequence (a) we hear two parallel melodies, each in a different frequency range. In Sequence (b) a single pitch is repetitively presented in the lower range, and this forms a ground against which the melody in the upper range is heard. Parts (a) and (b) illustrate the sequences in musical notation. The same sequences plotted as log frequency versus time are shown in (a') and (b'). [(a) From G. P. Telemann, *Capriccio for recorder and basso continuo*. (b) From G. P. Telemann, *Sonata in C major for recorder and basso continuo*. (a') and (b') From Musical illusions, D. Deutsch. Copyright © 1975 by *Scientific American*, Inc. All rights reserved.]

Dowling (1973b) performed an experiment to investigate the effect of frequency disparity on the perception of interleaved melodies. He presented the notes of two well-known melodies alternately at a rate of 8 sec^{-1}, and found that recognition of these melodies was very difficult when their pitch ranges overlapped. Recognition became increasingly easy as one of the melodies was transposed, so that the ranges of the two melodies gradually diverged.

Bregman and Campbell (1971) investigated this question from another point of view. They presented sequences of tones at a rate of 10 sec^{-1} that alternated between two widely different pitch ranges. They found that listeners had considerable difficulty in perceiving the order of tones in such sequences, though this problem did not arise when the tones were close in pitch. They concluded that the higher and lower tones formed separate perceptual streams, and that at rapid presentation rates order relationships cannot be formed across the two streams.

The principle of grouping tonal stimuli by frequency range is also apparent in a dichotic listening situation. Deutsch (1975a,d) presented subjects with a major scale that was played simultaneously in both ascending and descending form, switching from ear to ear in such a way that when a note from the ascending scale was in one ear, a note from the descending scale was in the other, and vice versa. Most subjects perceived this sequence as two separate melodic lines, a higher one and a lower one, moving in contrary motion. Furthermore, there was a strong tendency to hear all the higher tones as emanating from one earphone, and all the lower tones as emanating from the other. The remaining subjects perceived the higher melodic line, but heard little or nothing of the lower. So for all subjects this dichotic sequence was grouped perceptually by frequency range.

As a further manifestation of the law of proximity as applied to frequency range, there is a strong tendency for the frequency of occurrence of a melodic interval to be inversely correlated with its size. This has been found true in very different types of music, including Western cultivated music (Fucks, 1962; Ortmann, 1926), current popular music (Jeffries, 1974), and the music of various primitive cultures (Merriam, 1964). Figure 3 shows the relative occurrence of melodic intervals as a function of their size in these different types of music, and it can be seen that the plots exhibit a striking similarity.

B. Groupings Based on Timbre

Timbre is often used as a marker or carrier of sequential configurations (Erickson, 1975). In the music of composers such as Haydn, Mozart, and Beethoven, adjacent phrases are often played by different instruments, which helps to delimit the structure of the composition. Furthermore, over-

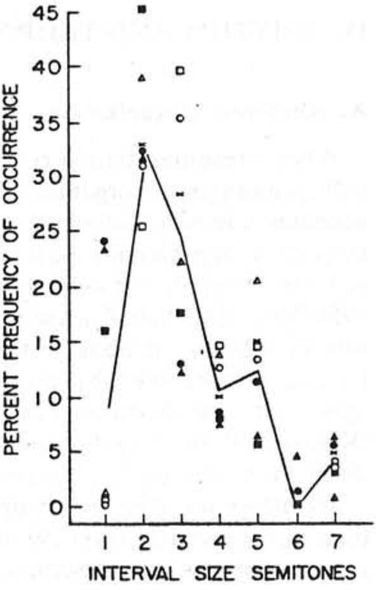

FIG. 3. Relative frequencies of occurrence of melodic intervals in the music of various cultures. Open circles represent Gêge (Dahomean-derived music of Brazil); open triangles represent Ketu (Yoruba-derived music of Brazil); open squares represent Rada (Dahomean-derived music of Trinidad); crosses represent Cheyenne Indian music; filled circles represent Romantic lieder; filled triangles represent a J. S. Bach orchestral part; and filled squares represent current popular songs. [Adapted in large part from Dowling (1967); data from Merriam (1964), Ortmann (1926), Fucks (1962), and Jeffries (1974).]

laps in pitch range between figure and ground are far more common when more than one instrument is involved. The formation of separate perceptual streams out of sounds of different timbre is also manifest in the finding that listeners have extreme difficulty in identifying the order of repetitive sequences of unrelated sounds (Warren, 1974; Warren & Obusek, 1972; Warren, Obusek, Farmer, & Warren, 1969).

C. Good Continuation

Divenyi and Hirsh (1974, 1975) provide evidence for the law of good continuation in the perception of rapid sequences of tones. They found that when subjects were required to identify the temporal order of three tones presented in rapid succession, identification was easier for sequences where the frequency change was unidirectional. The addition of a fourth tone generally had the effect of decreasing identifiability; however, performance was best when the fourth tone represented a continuation of a unidirectional frequency range. A related finding was obtained by Van Noorden (1975), who studied the minimum interval between successive components of three-tone sequences necessary to observe temporal coherence between the tones at various presentation rates. He found that for unidirectional three-tone sequences, temporal coherence was observed at rates of pitch change that were either equal to or higher than those necessary for two-tone sequences. However, for bidirectional three-tone sequences, the rate of pitch change had to be set considerably lower before coherence was perceived.

IV. RHYTHM AND TEMPO

A. Rhythmic Organization

When presented with a regular sequence of identical sounds, the listener will spontaneously organize these into subsequences, each consisting of an accented element followed by one or more unaccented elements. Such subjective organization occurs at presentation rates ranging from about 10 $\sec^{-1}$ to .5 $\sec^{-1}$, and appears to be optimal at rates of 2 or 3 $\sec^{-1}$ (Fraisse, 1956; Vos, 1973). As Fraisse points out, this correlates well with the distribution of melodic tempos in the cultivated music of our tradition. Indeed, tempo, as measured by the number of consecutive notes per unit time, appears to be distributed within this range in widely divergent cultures (Kolinski, 1959). The distributions of tempos in two cultures that are held to differ markedly on this measure are shown in Fig. 4.

Spontaneous rhythmic groupings are often formed simultaneously at more than one structural level (Woodrow, 1951; Vos, 1973). Thus, we may hear groups of four elements with the major accent on the first and a minor accent

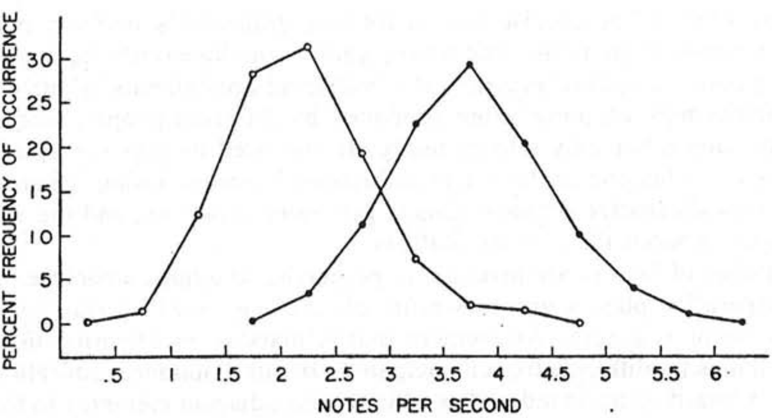

FIG. 4. Relative frequencies of occurrence of tempos in the songs of two cultures that diverge extremely in their average tempos. Open circles represent Dahomean songs, and filled circles represent North American Indian songs. Note that the shapes of the two distributions are remarkably similar and that the total range covered coincides with the range over which spontaneous rhythmic groupings are formed (see text). [Adapted from Kolinski (1959).]

on the third, or groups of eight elements with the major accent on the first and a minor accent on the fifth. This indicates that the mechanism underlying rhythmic grouping is hierarchical in nature. Further evidence has been provided in an experiment by Perkins (1974). He asked subjects to estimate the number of taps in sequences where the first of every 4 taps was stressed and the first of every 16 was doubly stressed. He found that errors differed from the correct responses more frequently by multiples of 4 and 16 than by adjacent numbers and concluded that the subjects were organizing the sequences hierarchically in accordance with the imposed accents. In another experiment, Sturges and Martin (1974) presented subjects with continuous sequences of 14 or 16 binary elements. These were either seven- or eight-element patterns repeated once, or patterns that were changed slightly on repetition; the subjects' task was to recognize sequences containing exact repetitions. They found that patterns that exhibited a simple hierarchical structure were better recognized than others and that eight-element patterns so structured were better recognized than seven-element patterns, even though they contained more items.

In actual music, it is clear that rhythm is hierarchically structured (Cooper & Meyer, 1960). However, such organization may be very complex. For instance, groupings at different levels often consist of different numbers of elements, and even at the same level groupings containing different numbers of elements may be presented in succession. Simultaneous polyrhythms also occur; however, these are very difficult for a single individual to generate. (The reader may try tapping a repetitive pattern consisting of three equally spaced beats with one hand and, simultaneously, four equally spaced beats with the other.) The precise reason for this difficulty is unclear, but the problem seems to lie in dividing a time period simultaneously into overlapping segments. In actual practice, the individual components of simultaneous polyrhythms are most often produced by different people, who need listen to each other only intermittently. It may well be that such music is perceived in a fashion analogous to ambiguous figures in vision, where only one of two alternative organizations is perceived at a time, and the viewer alternates between the two alternatives.

A number of factors are involved in perceiving rhythmic structure. Temporal separation plays a very powerful role (Handel, 1973; Restle, 1972). A second factor is accent. An element that is marked for attention in some way, such as by differing from its neighbors in pitch, loudness, duration, or timbre is heard as accented and combines with adjacent elements to form a group. (However, if an element differs too strongly from its neighbors, it dissociates itself instead [Miller & Heise, 1950].) Furthermore, a set of elements standing in a particular relationship will be grouped together, such as a run of identical elements (Garner, 1974) or similar elements. Groupings will also be formed depending on the abstract structure of the sequence (Simon & Sumner, 1968).

Rhythmic patterning has a powerful influence on perception of the tonal structure of a sequence. First, the formation of subsequences divides the information in manageable chunks, and so facilitates the perception of relationships within each chunk. In one experiment, Dowling (1973c) presented subjects with four five-note phrases that were separated by pauses. Subjects were then required to recognize a five-note test item that corresponded either to one of the phrases or to the last three notes of one phrase and the first two notes of the next. Recognition was better for items that corresponded to separate phrases than for items that crossed phrase boundaries, showing that the information was organized in accordance with the temporally determined phrase structure.

Second, such patterning is important in perceiving higher-order tonal relationships between subsequences. In music, a phrase is often repeated under some transformation with the rhythm intact. Here the corresponding elements of the two phrases stand in identical rhythmic relationship to their neighboring elements. This one-to-one correspondence of the rhythmic position of each element serves to enhance perception of the tonal correspondence. An example of such rhythmic "tagging" is shown in Fig. 5.

Accent provides another important function. Restle and Brown (1970) have shown that when a sequence of unaccented elements forms a structural tree, errors in learning increase monotonically with the level of transformation on the tree. Thus, in the absence of accent, higher-order relationships are perceived with difficulty. However, if the first element of each group is accented, perceptual relationships between these elements is enhanced, and this results in better perception of the higher-order structure. Indeed, a

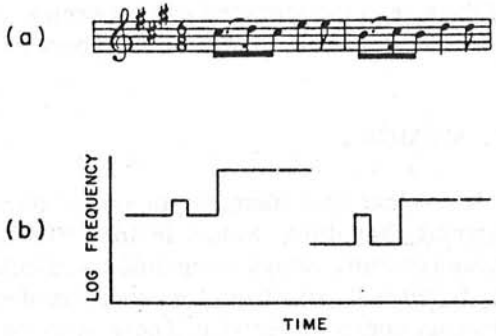

FIG. 5. In rhythmic tagging, the same melodic configuration is presented twice at different positions along the scale, with the same sequence of durations. Thus, each element is tagged by its rhythmic position in the sequence, and the corresponding elements of the two phrases are tagged identically. (a) The sequence in musical notation. (b) The sequence plotted as log frequency versus time. [From W. A. Mozart, *Sonata XVI in A major*.]

common musical technique is to present a sequence consisting of accented elements interspersed with unaccented elements. In this way the accented elements provide the figure and the unaccented elements provide the ground, yet unitary subgroups are still formed out of accented combined with unaccented elements.

B. Psychophysical Measurements

When listeners are required to make discrimination judgments involving time intervals between sound pairs, varying results are obtained depending on the stimulus conditions. Thus, with a standard interval that varied randomly between .63 and 640 msec, a just noticeable difference (jnd) of 40 msec was obtained at a standard of 100 msec (Abel, 1972). However, with the standard interval fixed at 100 msec, a jnd of 5 msec was obtained instead (Van Noorden, 1975). Michon (1964) required subjects to discriminate the tempos of sequences of regularly presented clicks. Plotting the results as Weber fractions, he found two maxima of sensitivity; one at an interval of 110 msec with $\Delta t/t$ of about .01, and the other at an interval of about 600 msec with $\Delta t/t$ of about .02.

A further interesting question concerns the detection of rhythmic irregularity in a sound sequence. Lunney (1974) asked subjects to adjust the temporal position of every fourth in a sequence of metronome pulses that were otherwise isochronous, so that a rhythmic irregularity was just detected. Repetition rates ranging from 30 to 3200 msec were explored, and sensitivity was found to decrease exponentially with the duration of the interval between repetitions. At a repetition rate of 100 msec, a jnd of 3–4 msec was obtained. Van Noorden (1975) studied the effect of frequency disparity on detection of rhythmic irregularity in a sequence of alternating tones. He found that sensitivity at high repetition rates was markedly reduced as the alternating tones diverged in frequency. This study may be related to an earlier finding by Divenyi and Hirsh (1972) that discrimination of the size of the temporal gap between a tone pair deteriorates with increasing pitch separation between members of the pair.

V. MEMORY

It is clear that memory for music must involve a large set of parallel systems that differ widely in their characteristics. Such systems include those retaining values along unidimensional acoustic continua, those retaining features formed from combinations of such values, and those involved in abstract cognitive activity. These systems must differ widely in the persistence with which they retain information. We know from general considerations that memory for certain relationships between values of an acoustic attribute persists much longer than memory for the absolute values themselves. Thus we can recognize a melody or a harmonic sequence when it is transposed to a different key, and, provided sufficient time has elapsed, we generally have difficulty in determining whether a key change has occurred. Here the absolute pitch information is lost, and recognition can be based only on the abstracted relationships (Attneave & Olson, 1971; Deutsch, 1969). Similarly, durational relationships are abstracted to produce rhythmic patterns, and these are easily recognized when played in a different tempo. However, certain complex acoustic stimuli deteriorate rapidly in memory (Guttman & Julesz, 1963; Pollack, 1972).

One question of importance concerns the nature of the influences acting on musical information in storage. This has been studied in detail in the case of tonal pitch. Deutsch (1970a) has shown that memory for pitch is not subject simply to general factors, such as attention distraction or capacity limitation, in a system that is either indifferent to the kind of information it holds or that retains any type of acoustic information. Subjects compared two tones for pitch when these were separated by a retention interval during which certain other acoustic stimuli were presented. It was found that a severe impairment of performance resulted from interpolating a sequence of six tones, even though the subjects were asked to ignore them. However, only a minimal impairment occurred when six spoken numbers were interpolated instead, even when the subjects were asked to recall the numbers in addition to making the pitch-comparison judgment.

Further studies have shown that the system retaining pitch information is precisely and systematically organized. Deutsch (1972b) required subjects to compare the pitch of two test tones that were separated by a sequence of six interpolated tones. In the second serial position of the intervening sequence was placed a tone whose pitch bore a critical relationship to the first test tone. This relationship varied in equal steps of one-sixth of a tone between identity and a whole-tone separation. It was found that when the critical interpolated tone was identical in pitch to the first test tone, memory facilitation occurred. Errors rose progressively as the pitch difference between these tones increased, peaked at a separation of two-thirds of a tone, and then returned to baseline. Thus, a specific disruptive effect was demonstrated, which varied as a function of the log-frequency difference between the first test tone and the critical interpolated tone.

This specific disruptive effect was further explored using critical interpolated tones that were a semitone removed from the first test tone. It was found that the inclusion in the intervening sequence of two critical tones, one a semitone higher than the first test tone and the other a semitone lower, produced a significantly greater increase in errors than the inclusion of only one critical tone. (Deutsch, 1973a). Furthermore, the size of this disruptive effect was found to be constant whether the critical intervening tone was placed in the second or in the fifth serial position of a sequence of six intervening tones (Deutsch, 1975b).

Given these findings, Deutsch and Feroe (1975) have put forward a specific model of the organization of pitch memory. They propose the existence of a memory array whose elements are activated by tones of specific pitch. These elements are spaced along the array so that elements activated by tones separated by the same distance in log-frequency units are separated by the same distance along the array. It was further assumed that these elements are linked as a recurrent lateral inhibitory network, analogous to those investigated in systems processing incoming sensory information. Evidence for lateral inhibitory interactions in the auditory system has been obtained at the psychophysical level by Carterette, Friedman, Lovell (1969) and Houtgast (1972). At the neurophysiological level, evidence for both peripherally and centrally acting lateral inhibition has also been found (Klinke, Boerger, & Gruber, 1969; Sachs & Kiang, 1968). Indeed, the relative frequency range over which centrally acting lateral inhibition has been plotted corresponds well with that for the present interactive effect.

This hypothesis was given further strong support from an experiment demonstrating disinhibition in pitch memory (Deutsch & Feroe, 1975). In this experiment, subjects were required to compare two tones for pitch when these were separated by a sequence of six intervening tones. In the second serial position of the intervening sequence was placed a tone that was two-thirds of a tone removed from the first test tone. Errors were then plotted as a function of the pitch of a further tone, placed in the fourth serial position, whose relationship to the tone in the second serial position varied in steps of one-sixth of a tone from identity to a whole-tone separation. As shown in Fig. 6, the curve produced was roughly the inverse of the curve plotting the original disruptive effect. Furthermore, in cases where the tone in the fourth serial position was two-thirds of a tone removed from the tone in the second serial position, errors were significantly lower than in the baseline condition, where the tone in the fourth serial position was outside the critical range. As a further test of the disinhibition hypothesis, a baseline

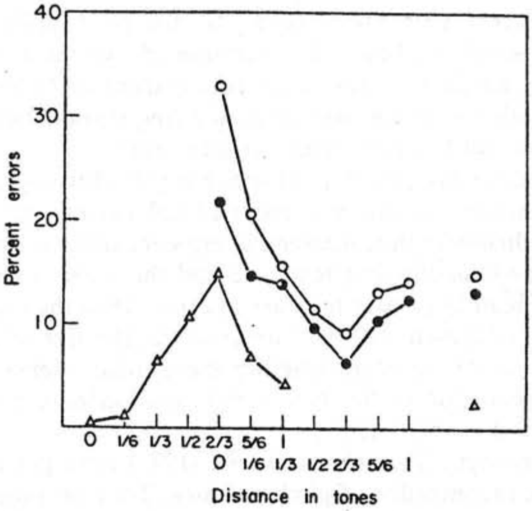

FIG. 6. Percent errors in pitch recognition as a function of various relationships within a tonal sequence. Subjects made pitch-comparison judgments between two tones that were separated by a sequence of interpolated tones. Open triangles display percent errors as a function of the pitch relationship between the first test tone and a critical interpolated tone. (Open triangle at right shows percent errors where no tones were interpolated within the critical range under study.) Filled circles display percent errors in the experiment where a tone that was two-thirds of a tone removed from the first test tone was always interpolated. Errors are plotted as a function of the pitch relationship between this tone and a second critical interpolated tone, which was farther removed along the pitch continuum. Open circles display percent errors for the same experimental conditions predicted theoretically. (Filled and open circles at right show percent errors obtained experimentally and assumed theoretically where no further critical tone was interpolated.) [From D. Deutsch & J. Feroe, Disinhibition in pitch memory, *Perception & Psychophysics*, 17, 320–324. Copyright 1975 by the Psychonomic Society, Inc. Reprinted by permission.]

curve for the first-order inhibitory effect was obtained using subjects selected on the same criterion as for the disinhibition experiment. These parameters were then used to plot the theoretical disinhibition function. As is also shown in Fig. 6, there was a very good correspondence between the theoretical function and that produced experimentally.

In sequences where the test tones differed in pitch, a further disruptive effect could also be demonstrated. If a tone of the same pitch as the second test tone was included in the intervening sequence, there resulted a substantial increase in errors of misrecognition (Deutsch, 1970b, 1972c). The error rate for such sequences was substantially higher than for sequences including a tone that was also a semitone removed from the first test tone but on the opposite side of the pitch continuum (Deutsch, 1973a). This misrecognition effect was also highly sensitive to the serial position of the repeated tone, being substantially more pronounced when this tone was placed in the second serial position of a sequence of six interpolated tones than when it was placed in the fifth serial position. It has been argued on these and other grounds that this disruptive effect is due to deterioration of information along a temporal, or order, continuum (Deutsch, 1972c, 1975b).

In a further study, specific disruptive effects in pitch memory were found to generalize from one octave to another (Deutsch, 1973b). The amount of generalization varied depending on the octave in which the critical interpolated tones were placed. The disruptive effects from tones that were displaced to a higher octave were greater than from tones displaced to a lower octave. From an analysis of the pattern of errors it was concluded that these disruptive effects took place along both a tone-height and a tone-chroma array.

Memory for pitch in a sequential setting is also subject to consolidation effects (Deutsch, 1975c). When the first test tone is repeated among the interpolated tones, there results a reduction in errors, even when compared with sequences containing a smaller number of interpolated tones. This improvement in performance is highly sensitive to the serial position of the repeated tone, being substantial and highly significant when this tone occurs early in the intervening sequence, but small and insignificant when it occurs late in the sequence.

When we consider absolute levels of performance, pitch memory in a sequential setting is found to be remarkably poor (Deutsch, 1970a; Pollack, 1952, 1964). This emphasizes the involvement of additional memory systems in processing tonal information. For instance, memory for abstracted tonal relationships plays an important role, and this has not been systematically studied as yet. Furthermore, the use of scales with strong a priori and transitional probabilities enables us to draw on information embedded in long-term memory in a short-term context also. Francès (1972) and Zenatti (1969) have shown that short-term recognition of melodies is superior for melodies that are in our tonal system than for those that violate this system. Furthermore, Miller and Cuddy (1972) have demonstrated that recognition of a tonal sequence that is preceded and followed by other tones is improved when the context tones establish a tonality. In music involving harmonic sequences, the strong transitional probabilities between root progressions further reduce the processing load. In addition, if music has a strong abstract structure (Restle, 1970; Simon & Sumner, 1968), the information can be retrieved by reference to this structure.

VI. TIMBRE PERCEPTION

The perception of timbre is a highly complex phenomenon that is, at present, little understood. Timbre is generally defined as that attribute by which two tones can be distinguished from each other when they are equal in pitch, loudness, and duration. Clearly, this definition is consistent with an enormous range of phenomena.

A. Steady-State Tones

Classically, investigations into timbre have been concerned with tones in the steady state. Helmholtz (1954) proposed that differences in the timbre of complex tones depended on the strength of their various harmonics. He concluded that simple tones sounded pleasant though dull at low frequencies, complex tones with moderately loud lower harmonics sounded richer but still pleasant, complex tones with strong upper harmonics sounded rough and sharp, and complex tones consisting only of odd harmonics sounded hollow. Helmholtz further asserted that the timbre of complex tones was affected little if at all by differences in phase relations between the partials. Plomp and Steeneken (1969) investigated this question and concluded that the maximum effect of phase on timbre was indeed very small; in fact, quantitatively smaller than the effect of changing the slope of the amplitude pattern by 2 dB per octave.

Other experiments by Plomp and his collaborators have demonstrated the importance of the critical band in timbre perception. Plomp (1964) and Plomp and Mimpen (1968) found that the number of partials that could be distinguished was limited to the lower five to seven harmonics, depending on the frequency of the fundamental. The higher harmonics, falling within the same critical band, could not be distinguished from each other. Furthermore, Plomp (1970) examined perceptual similarities between pairs of complex tones. For each tone the amplitude pattern of the waveform was quantified by taking the intensity levels for 18 consecutive one-third octave filters, corresponding to critical bands. Differences between the tone pairs were then computed by summing the differences in intensity levels for the 18 consecutive filters, so that the tones were placed as points in an 18-dimensional space. Using the multidimensional scaling technique, he found a good correlation between the perceptual similarities taken from subjects' ratings and the physical differences as determined by this method.

One issue of importance concerns whether timbral quality is based on the relationship between the frequency region of a formant and its fundamental or on the absolute level of the formant, irrespective of the fundamental frequency. Some musical instruments, such as the violin and the bassoon, have spectrum envelopes that are fixed in frequency (Fransson, 1966; Jansson, 1966). On the other hand, the spectral maxima of the flute change in proportion to changes in the frequency of the fundamental (Miller, 1916). Examination of the spectra of musical instruments therefore produces ambiguous expectations. Slawson (1968) asked subjects to make similarity judgments between pairs of complex tones whose fundamental frequencies, two lower formant frequencies, and higher formant frequencies were independently varied. With the fundamental frequency of the second tone of each pair an octave above the first, Slawson found timbral quality to be best preserved when the two lower formants were transposed by about 10% of the transposition of the fundamental. This result therefore supports a modified fixed-formant model of timbre perception. Furthermore, Plomp and Steeneken (1971) presented pulse trains through filters having different center frequencies. They found that tones filtered at fixed frequencies were judged as substantially more similar than tones filtered at frequencies relative to their pulse rates. These authors concluded that timbre was derived from the absolute position of the envelope of the frequency spectrum.

B. Tones Generated by Natural Instruments

With tones generated by natural instruments, temporal characteristics have been found to be very important determinants of timbre (Risset & Matthews, 1969). Traditionally, such tones have been held to consist of three temporal segments: (*a*) the attack, which includes the initial transient; (*b*) the steady state; and (*c*) the decay. Various studies have demonstrated that the attack segment is of particular importance (Berger, 1964; Goude, 1972; Grey, 1975; Saldanha & Corso, 1964; Wessel, 1974). Identification of the steady-state portion has been found to be superior for tones with vibrato (Saldanha & Corso, 1964). The decay portion appears to be of little importance (Saldanha & Corso, 1964). Risset (1966) made a careful study of the perceptual properties of trumpet tones, and concluded that three features were particularly significant: (*a*) the relationships between the attack times of the different harmonics, with successively higher harmonics taking longer to appear; (*b*) the low-amplitude, fast-frequency fluctuations; and (*c*) the harmonic content, including stronger high-frequency components with increases in overall intensity.

Interest has developed in plotting perceptual similarities between different instruments in a multidimensional timbral space. Carterette and Miller (1974) and Miller and Carterette (1975) have shown that such timbral spaces are related to musical training in a complex fashion. In an experiment where fundamental frequency was one of the dimensions varied, no differences were found between musical and nonmusical subjects, due to the overwhelming salience of this dimension. However, when fundamental frequency was held constant so that harmonic and envelope structure were used in judgments, differences between musical and nonmusical subjects emerged. Wessel (1973), studying tones of the same fundamental frequency and duration which were taken from nine orchestral instruments, concluded that these could be arranged along two perceptual dimensions. The first dimension related to the distribution of energy in the steady-state region of the spectrum. Tones with more high-frequency energy appeared at one end of the scale, and those with energy concentrated in the lower harmonics at the other end. The second dimension related to the onset patterns of the tones. Tones whose low-order harmonics entered more rapidly appeared at one end of the scale, and those whose high-order harmonics entered more rapidly appeared at the other. Grey (1975) studied 16 instrumental tones that were resynthesized by computer, and equated for pitch, loudness, and duration. He concluded that subjects' similarity ratings were consistent with a three-dimensional solution. The first dimension related to the spectral energy distribution of the tones: At one extreme, the tones had narrow bandwidths and a concentration of low-frequency energy; at the other extreme the tones had a wide bandwidth and less concentration of energy in the lower harmonics. The second dimension related to the distribution of energy in the attack segment: At one extreme, tones displayed high-frequency, low-amplitude energy (most often, inharmonic energy) in this segment. At the other extreme, there was no high-frequency precedent energy in the attack. The third dimension was given two alternative interpretations: It was considered either to relate to the form of onset–offset patterns of the tones, especially to the amount of synchronicity in the attacks and decays of the upper harmonics, or, alternatively, to represent a cognitive dimension along which stimuli were grouped according to instrumental family (strings, brass, and woodwinds).

In a recent set of experiments, Cutting and Rosner (1974) and Cutting, Rosner, and Foard (1976) have found that timbre based on rise time differences is categorically perceived. Sawtooth sounds of rapid onset sound as though coming from a plucked string, and those of slower onset from a bowed one. Discrimination of such sounds as "pluck" or "bow" was good across categorical boundaries, but poor within a category. Further, prolonged exposure to one such stimulus caused a shift in the rise time boundary between categories, analogous to shifts found in perception of speech sounds.

VII. MUSICAL PREFERENCES

A comprehensive discussion of musical preferences is outside the scope of the present chapter, and the reader is referred to Meyer (1956), Farnsworth (1958), Pratt (1968), and Berlyne (1971) for reviews of the question. Considerable discussion has centered on whether such preferences are innate or culturally determined, and it would appear that a position at either extreme is untenable. Farnsworth (1958) pointed out that many differences in the musical styles of different cultures. On the other hand, certain universal tendencies, such as a general range of tempo (see Fig. 4), a spectrum of probability of occurrence for different melodic intervals (see Fig. 3), octave duplications in harmony, and so on, have also been demonstrated. So it would seem that music is subject to universal laws, but is also subject to cultural variation within the framework imposed by these laws.

Interest has developed in the relationship of various stimulus parameters to aesthetic preference. Such studies have shown, as might be expected, that people tend to prefer intermediate values along a given stimulus dimension and reject the extremes. Vitz (1972) observed that preference for pure tones varied as a U-shaped function of frequency, with the most preferred tones ranging from 400 to 750 Hz. A similar function was obtained for tonal intensity, with the most preferred intensity at around 50 dB. We would also expect, as argued by Attneave (1959), that aesthetic preferences for tonal sequences differing in amount of uncertainty or complexity would be optimal at some level intermediate between homogeneity and chaos. Indeed, such U-shaped functions have been obtained (Berlyne, 1971). Interestingly, such preferences have also been shown to interact with the emotional state of the listener. In an experiment by Konecni, Crozier, and Doob (1976) subjects were made angry by a confederate, and then were given either a chance to administer what they thought to be electric shocks to this person, or were given no such opportunity. A control group of subjects received neither treatment. All subjects then listened to tonal sequences and chose between sequences of differing complexity. It was found that the angered subjects who had no opportunity to retaliate chose the simple sequences more often than either the controls or the angered subjects who had been seemingly able to hurt their annoyer. There was no significant difference in preference between the latter two groups. In a further experiment, Konecni (1975) demonstrated that listening to such sequences when emotionally aroused had specific effects on subsequent social behavior. Subjects were either angered by a confederate or were treated neutrally. Some subjects then listened to tonal sequences that varied in loudness and complexity, and a control group of subjects listened to no sequences. All subjects were then given an opportunity to administer apparent electric shocks to a confederate. It was found that the behavior of the angered subjects was affected by the nature of the sequences they had heard. Those exposed to loud or complex sequences administered considerably more shocks than those exposed to no sequences; however, those exposed to simple and soft sequences administered fewer shocks than the controls. In contrast, there was no effect of the sequences on the behavior of subjects who had not been angered. These two experiments demonstrate an important link between aesthetic preference and social interaction.

VIII. NEUROLOGICAL SUBSTRATES

The neurological substrates of music are, at present, poorly defined. We know, however, that they are not identical with those of speech. In Broca's aphasia, verbal output is grossly impaired, yet the patient can generally sing without difficulty (Geschwind, 1969). Indeed, there are several reports of professional musicians whose work remained outstanding in spite of severe aphasia. Critchley (1953) mentioned the case of a severely aphasic patient who conducted his own orchestra. Luria, Tsvetkova, and Futer (1965) described a distinguished composer who, after being struck with severe aphasia, continued to produce compositions that won him the highest acclaim. If we also consider that singing ability may be lost without concomitant aphasia (Brain, 1965), we must conclude that speech and music involve specialized brain regions. This does not, however, imply a total dissociation between the substrates for these two functions, which are both highly complex and may well share common elements. The evidence simply demonstrates that there are unique regions essential to each.

Most authors agree that the temporal lobes are involved in music perception. Lesions of the temporal lobe, particularly of the anterior temporal region, have been found to produce sensory amusia (Henschen, 1926; Luria, 1966; Milner, 1962; Wertheim, 1963). Furthermore, reports of sounds such as buzzing, humming, and even complex music are evoked by temporal-lobe stimulation (Penfield & Rasmussen, 1968). In contrast to music perception, musical expression involves the motor system, and expressive amusia has been found to result from damage to a variety of brain regions (Luria, 1966; Wertheim, 1963).

Considerable interest has focused on the question of hemispheric specialization and musical function, and here again it is necessary to distinguish between musical expression and musical perception.

Musical expression, as opposed to speech, does not require the presence of the dominant hemisphere. After dominant hemispherectomy, patients have been found to retain their singing ability (Gott, 1973; Smith, 1966). However, musical expression does not require the presence of the nondominant hemisphere, either, as patients with their nondominant hemisphere removed can also sing without difficulty (Smith, personal communication). In one study, Gorden and Bogen (1974) compared the relative roles of the two hemispheres in singing, using the technique of intracarotid injection of sodium amylobarbitone (which produces a transient inactivation of the hemisphere on the side of the injection). Injection in the right carotid artery produced marked deficits in singing, although speech remained relatively intact. After injection in the left carotid, singing was less disturbed than speech. The authors argue from these findings that whereas speech production is overwhelmingly a function of the dominant hemisphere, the nondominant hemisphere plays a more prominent role in singing. It was noted, however, that after injection in the right carotid the patients recognized their own singing efforts as poor and were able to recognize songs sung by the examiner. Inactivation of the nondominant hemisphere did not, therefore, appear to produce deficits in music perception. The relationship of these findings to the findings resulting from hemispherectomy still needs to be explored.

The evidence on hemispheric specialization in music perception is quite complex. From observations on patients with sensory amusia, the dominant hemisphere has generally been implicated to a larger extent than the nondominant (Henschen, 1926; Kleist, 1962; Wertheim, 1963), though lesions of the nondominant hemisphere have also been found to produce impairment (Bogen, 1969; Luria, 1966). Studies involving specific musical attributes have also produced varying results. Thus Milner (1962) administered the Seashore Tests of Musical Talents to patients before and after unilateral temporal lobectomy. Removal of the anterior temporal region of the nondominant side produced significant deficits in the ability to discriminate tonal quality and in the ability to identify which member of a tonal sequence was altered on second presentation. These deficits were not found with lesions of the dominant side. Furthermore, Shankweiler (1966) found that the ability to recognize 4-sec excerpts from instrumental chamber music was selectively impaired by lesions of the right temporal lobe, but not of the left temporal lobe. Vignolo (1969) has reported that patients with lesions of the nondominant hemisphere showed deficits in discriminating sound quality. On the other hand, lesions of the dominant hemisphere (rather than the nondominant) have been found to produce deficits in tasks involving judgments of temporal order (Carmon & Nachshon, 1971; Efron, 1963; Swisher & Hirsh, 1972) and rhythm (Luria, 1966; Subzinski, 1969).

These findings indicate that different musical attributes might be handled predominantly in different hemispheres. In particular, it seems that the nondominant hemisphere predominates in processing the quality of musical sounds, whereas the dominant hemisphere predominates in processing sequential relationships. This conclusion receives further support from dichotic listening studies. A left-ear advantage has been noted in right-handed persons with respect to recognizing auditory materials of complex sound quality, such as melodies generated by musical instruments (Kimura, 1964) or by humming (King & Kimura, 1972), sonar signals (Chaney & Webster, 1966), and environmental sounds (Curry, 1967; Knox & Kimura, 1970). Gordon (1970) failed to obtain evidence of a left-ear advantage in recognizing dichotically presented melodies played on a recorder, but did find a left-ear advantage in recognizing dichotically presented chords that were generated by an electronic organ. In the chords test, two simultaneous tone pairs were presented, one to each ear; therefore, the auditory information present at any one time was rich in timbral and simultaneous interval information.

Other dichotic listening studies have detected a right-ear advantage in processing sound sequences in which evaluation of tonal quality or simultaneous relationships was not involved. Halperin, Nachshon, and Carmon (1973) presented subjects with dichotic sequences varying in frequency or duration. They found that as the number of frequency or duration transitions increased from zero to two, superiority of performance shifted from the left to the right ear. Also, Deutsch (1974, 1975d) presented subjects with a dichotic sequence of sine wave tones, which alternated in pitch from one octave to another such that when the right ear received the high tone the left ear received the low tone, and vice versa. Righthanders here tended to perceive the pitch information presented to the right ear rather than to the left. Furthermore, Robinson and Solomon (1974) presented subjects with dichotic rhythmic stimuli, produced by pure tones of invariant pitch. They also obtained results indicating a right-ear advantage in recognizing these sequences. Also, Papçun, Krashen, Terbeek, Remington, and Harshman (1974) explored the effect of dichotic presentation on Morse code signals. They found that experienced Morse code operators showed a right-ear superiority in perceiving these sequences. Naive subjects also showed a right-ear superiority, provided that the patterns were restricted to seven or fewer elements. With patterns of more than seven elements, ear superiority shifted from right to left. Papçun *et al.* hypothesized that when the subjects were able to deal with the individual elements and their sequential relationships, then processing took place in the dominant hemisphere. However, when the stimuli became too complex, subjects were forced to adopt a holistic strategy instead, shifting the processing from the dominant hemisphere to the nondominant. A similar line of reasoning was advanced by Bever and Chiarello (1974), who found a right-ear advantage among trained musicians in the recognition of monaurally presented tonal sequences. However, nonmusicians, who showed inferior performance, were found to have a left-ear advantage. The authors concluded that the musicians were processing the sequences analytically—the nonmusicians, holistically. Although the nature of the holistic processing underlying the left-ear advantage in the naive subjects of the experiments conducted by Papçun *et al.* (1974) and Bever and Chiarello (1974) remains to be clarified, these findings emphasize the specialization of each hemisphere in different aspects of musical processing.

References

Abel, S. M. Discrimination of temporal gaps. *Journal of the Acoustical Society of America*, 1972, **52**, 519–524.

Attneave, F. Stochastic compositional processes. *Journal of Aesthetics and Art Criticism*, 1959, **17**, 503–510.

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

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. P. 69.

Bake, A. *The music of India.* In E. Wellez (Ed.), *Ancient and oriental music.* London: Oxford Univ. Press, 1957. Pp. 195–227.

Berger, K. W. Some factors in the recognition of timbre. *Journal of the Acoustical Society of America,* 1964, **36,** 1888–1891.

Berlyne, D. E. *Aesthetics and psychobiology.* New York: Appleton, 1971.

Bever, T. G., & Chiarello, R. J. Cerebral dominance in musicians and nonmusicians. *Science,* 1974, **185,** 537–539.

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

Bogen, J. E. The other side of the brain II: An appositional mind. *Bulletin of the Los Angeles Neurological Society,* 1969, **34,** 135–162.

Boomsliter, P., & Creel, W. Extended reference: An unrecognized dynamic in melody. *Journal of Music Theory,* 1963, **7,** 2–22.

Brain, W. R. *Speech disorders* (2nd ed.). London: Butterworth, 1965.

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

Burns, E. M. Octave adjustment by non-western musicians. *Journal of the Acoustical Society of America,* 1974, **56,** S25–S26. (a)

Burns, E. M. In search of the shruti. *Journal of the Acoustical Society of America,* 1974, **56,** S26. (b)

Burns, E. M., & Ward, W. D. Categorical perception of musical intervals. *Journal of the Acoustical Society of America,* 1973, **54,** S96.

Carmon, A., & Nachshon, I. Effect of unilateral brain damage on perception of temporal order. *Cortex,* 1971, **7,** 410–418.

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

Chaney, R. B., & Webster, J. C. Information in certain multidimensional sounds. *Journal of the Acoustical Society of America,* 1966, **40,** 455–477.

Cooper, G. W., & Meyer, L. B. *The rhythmic structure of music.* Chicago: Univ. of Chicago Press, 1960.

Critchley, M. *The parietal lobes.* London: Arnold, 1953.

Curry, F. K. W. A comparison of left-handed and right-handed subjects in verbal and nonverbal dichotic listening tasks. *Cortex,* 1967, **3,** 343–352.

Cutting, J. E., & Rosner, B. S. Categories and boundaries in speech and music. *Perception & Psychophysics,* 1974, **16,** 564–570.

Cutting, J. E., Rosner, B. S., & Foard, C. F. Perceptual categories for musiclike sounds: Implications for theories of speech perception. *Quarterly Journal of Experimental Psychology,* 1976, **28,** 361–378.

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

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

Deutsch, D. Dislocation of tones in a musical sequence: A memory illusion. *Nature,* 1970, **5242,** 226. (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 Experinental Psychology,* 1972, **93,** 156–162. (c)

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

Deutsch, D. Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics,* 1973, **13,** 271–275. (b)

Deutsch, D. An auditory illusion. *Nature,* 1974, **251,** 307–309.

Deutsch, D. Two-channel listening to musical scales. *Journal of the Acoustical Society of America,* 1975, **57,** 1156–1160. (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. Musical Illusions. *Scientific American,* 1975, **233,** 92–104. (d)

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

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

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.

Divenyi, P. L., & Hirsh, I. J. Discrimination of the silent gap in two-tone sequences of different frequencies. *Journal of the Acoustical Society of America,* 1972, **52,** 166 (Appendix).

Divenyi, P. L., & Hirsh, I. J. Identification of temporal order in three-tone sequences. *Journal of the Acoustical Society of America,* 1974, **56,** 144–151.

Divenyi, P. L., & Hirsh, I. J. The effect of blanking on the identification of temporal order in three-tone sequences. *Perception & Psychophysics,* 1975, **17,** 246–252.

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

Dowling, W. J. Recognition of melodic transformations: Inversion, retrograde, and retrograde-inversion. *Perception & Psychophysics,* 1972, **12,** 417–421.

Dowling, W. J. The 1215-cent octave: Convergence of Western and Nonwestern data on pitch scaling. *Journal of the Acoustical Society of America,* 1973, **53,** 373. (a)

Dowling, W. J. The preception of interleaved melodies. *Cognitive Psychology,* 1973, **5,** 322–337. (b)

Dowling, W. J. Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics,* 1973, **14,** 37–40. (c)

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. Also, Abh. d. kgl. sachs. Ges. d. Wiss. math.-phys. Cl.b.II. 1855, 35.

Efron, R. Temporal perception, aphasia, and deja vu. *Brain,* 1963, **86,** 403–424.

Erickson, R. *Sound structure in music.* Berkeley: Univ. of California Press, 1975.

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

Farnsworth, P. R. *The social psychology of music.* New York: Holt, 1958.

Fraisse, P. *Les structures rhythmiques.* Paris: Erasme, 1956.

Francès, R. *La perception de la musique.* Paris: Vrin, 1972.

Fransson, F. The source spectrum of double-reed woodwind instruments. *Royal Institute of Technology, Stockholm, Speech Transmission Lab, QPSR,* 1966, **4,** 35.

Fucks, W. Mathematical analysis of the formal structure of music. *IRE Trans. Information Theory,* 1962, **8,** 225–228.

Garner, W. R. *The processing of information and structure.* New York: Wiley, 1974.

Geschwind, N. Problems in the anatomical understanding of the aphasias. In A. L. Benton (Ed.), *Contributions to clinical neuropsychology.* Chicago: Aldine, 1969.

Gordon, H. W. Hemispheric asymmetrics in the perception of musical chords. *Cortex*, 1970, **6**, 387–398.

Gordon, H. W., & Bogen, J. E. Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *Journal of Neurology, Neurosurgery, and Psychiatry*, 1974, **37**, 27–738.

Gott, P. Language after dominant hemispherectomy. *Journal of Neurology, Neurosurgery, and Psychiatry*, 1973, **36**, 1082–1088.

Grey, J. M. *An exploration of musical timbre*. Unpublished doctoral dissertation, Stanford Univ., 1975.

Guttman, N., & Julesz, B. Lower limits of auditory periodicity analysis. *Journal of the Acoustical Society of America*, 1963, **35**, 610.

Halperin, Y., Nachshon, I., & Carmon, A. Shift of ear superiority in dichotic listening to temporally patterned nonverbal stimuli. *Journal of the Acoustical Society of America*, 1973, **53**, 46–50.

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

Heise, G. A., & Miller, G. A. An experimental study of auditory patterns. *American Journal of Psychology*, 1951, **64**, 68–77.

Helmholtz, H. L. F. *On the sensations of tone, as a physiological basis of music*. New York: Dover, 1954. (Originally published 1862.)

Henschen, S. E. On the function of the right hemisphere of the brain in relation to the left in speech, music and calculation. *Brain*, 1926, **49**, 110–123.

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

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

Jansson, E. Analogies between bowed string instruments and the human voice, source-filter models. *Royal Institute of Technology, Stockholm, Speech Transmission Lab, QPSR*, 1966, **3**, 4.

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

Kimura, D. Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, 1964, **16**, 355–358.

King, F. L., & Kimura, D. Left-ear superiority in dichotic perception of vocal nonverbal sounds. *Canadian Journal of Psychology*, 1972, **26**, 111–116.

Kleist, K. *Sensory aphasia and amusia*. The Myelo-Architectomic basis. Oxford: Pergamon Press, 1962.

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.

Knox, C., & Kimura, D. Cerebral processing of nonverbal sounds in boys and girls. *Neuropsychologia*, 1970, **8**, 117–137.

Kolinski, M. The evaluation of tempo. *Ethnomusicology*, 1959, **3**, 45–57.

Konecni, V. J. The mediation of aggressive behavior: Arousal level vs. anger and cognitive labeling. *Journal of Personality and Social Psychology*, 1975, **32**, 706–712.

Konecni, V. J., Crozier, J. B., & Doob, A. N. Anger and expression of aggression: Effects on aesthetic preference. *Scientific Aesthetics*, 1976, **1**, 47–55.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. Perception of the Speech Code. *Psychological Review*, 1967, **74**, 431–461.

Lunney, H. W. M. Time as heard in speech and music. *Nature*, 1974, **249**, 592.

Luria, A. R. *Higher cortical functions in man*. New York: Basic Books, 1966.

Luria, A. R., Tsvetkova, L. S., & Futer, D. S. Aphasia in a composer. *Journal of the Neurological Sciences*, 1965, **2**, 288–292.

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

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

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

Michon, J. A. Studies on subjective duration 1. Differential sensitivity on the perception of repeated temporal intervals. *Acta Psychologica*, 1964, **22**, 441–450.

Miller, D. C. *The science of musical sounds*. New York: Macmillan, 1916.

Miller, G. A., & Heise, G. A. The trill threshold. *Journal of the Acoustical Society of America*, 1950, **22**, 637–638.

Miller, J., & Cuddy, L. L. Tonality as a cue for melody recognition. *Experimental Report 72-1*, Queen's University, Kingston, Ontario, 1972.

Miller, J. R., & Carterette, E. C. Perceptual space for musical structures. *Journal of the Acoustical Society of America*, 1975, **58**, 711–720.

Milner, B. Laterality effects in audition. In V. B. Mountcatle (Ed.), *Interhemisphoric relations and cerebral dominance*. Baltimore: Johns Hopkins Univ. Press, 1962. Pp. 177–195.

Nettl, B. *Music in primitive culture*. Cambridge, Massachusetts: Harvard Univ. Press, 1956. (a)

Nettl, B. Unifying factors in folk and primitive music. *Journal of the American Musicological Society*, 1956, **9**, 196–201. (b)

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

Papçun, G., Krashen, S., Terbeek, D., Remington, R., & Harshman, R. Is the left hemisphere specialized for speech, language and/or something else? *Journal of the Acoustical Society of America*, 1974, **55**, 319–327.

Penfield, W., & Rasmussen, T. *The cerebral cortex of man*. New York: Hafner, 1968.

Perkins, D. N. Coding position in a sequence by rhythmic grouping. *Memory & Cognition*, 1974, **2**, 219–223.

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

Piston, W. *Counterpoint*. London: Gollancz, 1949.

Plomp, R. The ear as frequency analyzer. *Journal of the Acoustical Society of America*, 1964, **36**, 1628–1636.

Plomp, R. Timbre as a multidimensional attribute of complex tones. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Sijthoff: Leiden, 1970.

Plomp, R., & Mimpen, A. M. The ear as frequency analyzer II. *Journal of the Acoustical Society of Amerlca*, 1968, **43**, 764–767.

Plomp, R., & Steeneken, H. J. M. Effects of phase on the timbre of complex tones. *Journal of the Acoustical Society of America*, 1969, **46**, 409–421.

Plomp, R., & Steeneken, H. J. M. Pitch versus timbre. Paper presented at the Seventh International Congress on Acoustics, Budapest, 1971.

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

Pollack, I. The information of elementary auditory displays. *Journal of the Acoustical Society of America*, 1952, **24**, 745–749.

Pollack, I. Ohm's acoustical law and short term auditory memory. *Journal of the Acoustical Society of America*, 1964, **36**, 2340–2345.

Pollack, I. Memory for auditory waveform. *Journal of the Acoustical Society of America*, 1972, **51**, 1209–1215.

Pratt, C. C. *The meaning of music*. New York: Johnson Reprint Corp., 1968.

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. Serial pattern learning: Higher order transitions. *Journal of Experimental Psychology*, 1973, **99**, 61–69.

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

Revesz, G. *Zur Grundleguncy der Tonpsychologie*. Leipzig: Feit, 1913.

Risset, J. C. *Computer study of trumpet tones*. Murray Hill, New Jersey: Bell Telephone Laboratories, 1966.

Risset, J. C., & Matthews, M. V. Analysis of musical instrument tones. *Physics Today*, 1969, **22**, 23–30.

Robinson, G. M., & Solomon, D. J. Rhythm is processed by the speech hemisphere. *Journal of Experimental Psychology*, 1974, **102**, 508–511.

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

Sachs, M. B., & Kiang, N. Y-S Two-tone inhibition in auditory nerve fibers. *Journal of the Acoustical Society of America*, 1968, **43**, 1120–1128.

Saldanha, E. L., & Corso, J. F. Timbre cues for the recognition of musical instruments. *Journal of the Acoustical Society of America*, 1964, **36**, 2021–2026.

Shankweiler, D. Effects of temporal-lobe damage in perception of dichotically presented melodies. *Journal of Comparative and Physiological Psychology*, 1966, **62**, 115–119.

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

Siegel, W., & Sopo, R. Tonal intervals are perceived categorically by musicians with relative pitch. *Journal of the Acoustical Society of America*, 1975, **57**, S11.

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

Slawson, A. W. Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency. *Journal of the Acoustical Society of America*, 1968, **43**, 87–101.

Smith, A. Speech and other functions after left (dominant) hemispherectomy. *Journal of Neurology, Neurosurgery, and Psychiatry*, 1966, **29**, 467–471.

Stumpf, C., & Meyer, M. Maassbestimmungen uber die Reinheit consonanter Intervalle. *Beitr Akust. Musikwiss*, 1898, **2**, 84–167.

Sturges, P. T., & Martin, J. G. Rhythmic structure in auditory temporal pattern perception and immediate memory. *Journal of Experimental Psychology*, 1974, **102**, 377–383.

Subczinski, J. Described in R. Masland, "Brain Mechanisms Underlying the Language Function." In *Human communication and its disorders*. Bethesda, Maryland: National Institute of Neurological Diseases and Stroke, 1969.

Sundberg, J. E. F., & Lindquist, J. Musical octaves and pitch. *Journal of the Acoustical Society of America*, 1973, **54**, 922–929.

Swisher, L., & Hirsh, I. J. Brain damage and the ordering of two temporally successive stimuli. *Neuropsychologia*, 1972, **10**, 137–152.

Terhardt, E. Pitch shifts of harmonics, an explanation of the octave enlargement phenomenon. *Proceedings of the Seventh International Congress on Acoustics*, Budapest, 1971, 621–624.

Tovey, D. F. *The forms of music*. London: Oxford Univ. Press, 1957.

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

Vignolo, L. A. Auditory agnosia: A review and report of recent evidence. In A. L. Benton (Ed.), *Contributions to clinical neuropsychology*. Chicago: Aldine, 1969. Pp. 172–231.

Vitz, P. C. Preference for tones as a function of frequency (hertz) and intensity (decibels). *Perception & Psychophysics*, 1972, **11**, 84–88.

Vos, P. G. *Perception of metrical tone sequences*. Unpublished doctoral dissertation, University of Nijmegen, 1973. The Netherlands.

Ward, W. L. Subjective musical pitch. *Journal of the Acoustical Society of America*, 1954, **26**, 369–380.

Ward, W. D. Musical perception. In J. V. Tobias (Ed.), *Foundations of modern auditory theory* (Vol. I). New York: Academic Press, 1970. Pp. 407–447.

Warren, R. M. Auditory temporal discrimination by trained listeners. *Cognitive Psychology*, 1974, **6**, 237–256.

Warren, R. M., & Obusek, C. J. Identification of temporal order within auditory sequences. *Perception & Psychophysics*, 1972, **12**, 86–90.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. Auditory sequence: Confusions of patterns other than speech or music. *Science*, 1969, **164**, 586–587.

Wedin, L., & Goude, G. Dimension analysis of the perception of instrumental timbre. *Scandinavian Journal of Psychology*, 1972, **13**, 228–240.

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

Wertheim, N. Disturbances of the musical functions. In L. Halpern (Ed.), *Problems of dynamic neurology*. New York: Grune & Stratton, 1963.

Wertheimer, M. Untersuchungen sur Lehre von der Gestalt, II, *Psychologische Forschung*, 1923, **4**, 301–350.

Wessel, D. L. Psychoacoustics and music. *Bulletin of the Computer Arts Society*, 1973, **1**, 30–31.

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

Winograd, T. Linguistics and the computer analysis of tonal harmony. *Journal of Music Theory*, 1968, **12**, 2–49.

Woodrow, H. Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1224–1236.

Zenatti, A. Le developpement genetique de la perception musicale. *Monographies Francaises de Psychologie*, 1969, **17**, 1–110.