The representation of harmonic structure in music: Hierarchies of stability as a function of context

JAMSHED BHARUCHA*
Harvard University

CAROL L. KRUMHANSL
Cornell University

Abstract

The ability to appreciate most Western music presupposes cognitive structures capable of abstracting an underlying harmonic structure from a complex string of musical events. In this paper we provide a description of the listener's knowledge of hierarchies of harmonic stability. The organization of harmonic information may be summarized by six principles. Three of these principles—Key Membership, Intrakey Distance and Intrakey Asymmetry—govern harmonic organization independent of context. Three principles—Contextual Identity, Contextual Distance and Contextual Asymmetry—govern harmonic organization in the presence of a tonal context. Chords that are members of the same key are represented in a hierarchy of stability that is independent of context. Chords from different keys are represented in a hierarchy of stability that is dependent upon the prevailing context. Two different experimental tasks were used to provide convergent evidence for these principles: 1) multidimensional scaling of chords in the absence of any context or in the presence of different tonal contexts, and 2) recognition memory for chords in random and tonal contexts. It is suggested that harmonically stable chords function as cognitive reference points for the system as a whole. The importance of representations of hierarchies of harmonic stability is discussed with respect to generative accounts of musical competence.

*This research was supported in part by a grant from the National Science Foundation (BNS-81-03570) to Carol L. Krumhansl. The authors are grateful to David M. Green for the use of the Psychophysics Laboratory at Harvard and to William K. Estes for helpful suggestions. Murray Spiegel, David Wilson, Michael Hacker, Robert Nosofsky and an anonymous reviewer gave valuable advice. Mary Castellano assisted in running subjects and in data analysis. Reprint requests may be addressed to Jamshed Bharucha (who is now at Cornell) or Carol L. Krumhansl at Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853, U.S.A.
The perceptual processing of music involves the abstraction of underlying invariants in auditory events of considerable surface complexity. A piece of music can undergo substantial surface alteration (in terms of instrumentation, embellishment, transposition, or variation) and yet be recognized as the same piece. Only a level of description more abstract than that of the acoustic signal can capture the structural properties of music that enable the listener to perceive musical sequences with marked surface differences as similar. This ability suggests that the mental representation of music is governed by a system of abstract structural principles.

Substantial work has been done on the cognitive systems that enable us to process melodies, that is, sequences of tones (for example, Bartlett and Dowling, 1980; Cuddy et al., 1979; Deutsch, 1969, 1980; Deutsch and Feroe, 1981; Dewar et al., 1977; Dowling, 1978; Idson and Massaro, 1978; Krumhansl, 1979; Krumhansl and Shepard, 1979; Massaro et al., 1980). Only recently has experimental research been done on the cognitive systems that enable us to process harmonic organization in sequences of chords (Krumhansl, Bharucha and Kessler, 1982). In this paper, we argue that the extraction of harmonic structure is fundamental to the processing of Western music, and propose six basic structural principles governing the representation of harmony.

Music theoretic descriptions of musical structure

Music theorists and composers have been concerned with the formulation of theories of the abstract structure of music (Schenker, 1935/1979, 1906/1954; Schoenberg, 1969). The prevailing theory that has emerged offers a description of music in terms of two fundamental concepts, which are constitutive of musical structure referred to as tonal. The first is the hierarchy of stability that applies to the twelve tones of the chromatic scale within an octave range. The concept of stability refers to the fact that certain tones are perceived as more final and serve as better completions of melodic phrases than do others, which demand ‘resolution’ to more stable tones in the system. The more stable tones are generally those that appear more frequently, in prominent positions, and with rhythmic stress. In the hierarchy of stability, one single tone, called the tonic, is the most stable tone. It is also the first tone of the scales of Western tonal music. Following the tonic, the other tones of the scale are the next most stable, particularly the fifth and the third tones of the scale. The tones not contained in the scale are the least stable within the system. In Western music, the major and minor diatonic scales are most frequently employed. In Eastern music,
other scales are often used, but the hierarchies are often analogous. Since any tone can function as the tonic, in other words, melodies are transposable, the placement of the tonal hierarchy is dependent upon the context and is determined by the use of tones from a particular scale and their temporal ordering. In its most general form, this account seems to capture the basic structural organization of most music, Western and Eastern alike. Additional structure over and above this may or may not emerge cross-culturally, but this question is beyond the scope of this paper.

In Western music, a second fundamental organizing concept is the harmonic functioning of simultaneously sounded tones in chords. By far the majority of the work of music theorists has dealt with harmonic organization, which provides a description of the abstract structure of most Western music over the last four centuries. It is noteworthy that although many surface aspects of music have changed drastically over this time, the underlying harmonic structure has remained constant for the most part. Popular forms of music still cling to the traditional emphasis on triads built on the tonic (I), the dominant (V), and the subdominant (IV) chords.

The units of harmonic analysis, as suggested by music theory, are chords (or more precisely, root functions). The chords into which a piece of music is analyzed may not actually be present in the music as simultaneously sounded tones. They may, instead, be implied by a sequence of successive tones (i.e., a melody). The analysis of tonal melodies in music theory thus invariably makes reference to the underlying—or implied—harmonic structure. The music theoretic analysis of music into abstract chord functions suggests the possibility that listeners have an internal representation of these chord functions which may be activated not only by direct sounding of these chords, but also by certain melodic patterns.

The rhythmic analysis of a piece of music may also rely upon harmonic structure. It is often the case that the ‘strong beat’ may be found by the listener with no difficulty even though it may be cued neither by loudness nor temporal parameters. In such cases, the strong beat will be perceived at the position of chords that are harmonically stable in the prevailing context. The stress is therefore implied by harmonic functions even though an actual acoustic stress may occur in a different temporal position in the sequence or may not occur at all.

It is of course true that just as melodic and rhythmic analyses rely on harmonic structure, the converse holds as well to some extent. For instance, harmonic analysis often involves melodic considerations such as voice-leading (which governs the expectations of melodic progression), and the expectations for harmonic progression may be strongly influenced by metrical or rhythmic factors. However, it is harmonic structure that music
theorists have been predominantly concerned with, and with which they have been remarkably successful in constructing sophisticated analytic theories. Although there is by no means unanimous agreement on the fine details of a standard theory of harmony, the principles are codified in rigorous textbooks (e.g., Piston, 1962) and are routinely taught in courses on harmony.

In sharp contrast, a comprehensive theory of melodic structure has eluded music theorists. Rarely does one read a textbook or take a course on melody. This is not because melodic structure is necessarily simpler than harmonic structure. Rather, it is because, with the exception of principles of voice-leading, only vague principles of good melodic form have been forthcoming. Furthermore, as stated earlier, attempts at rigorous analyses of melodic structure invariably invoke the harmonic structure that is implied by the melody.

The analysis of root functions in music theory is the identification of the stable tones of the composition as components of chords that function within the prevailing tonality or key. Seven chords constitute the basic set of chords drawn from any given key, which are those built on the seven steps of the diatonic scale. Often the Roman numerals I–VII are used to indicate the scale step on which a chord is built (i.e., the root tone). For the most part, chords of three tones (triads) are used, with a second tone sounded as either a minor or major third (three or four chromatic steps, respectively) above the root tone, and a third tone sounded as either a major or minor third above the second tone. Depending on the particular combination of major or minor thirds, the resulting chord is major, minor, diminished, or augmented. The chord type is determined by the position of the root of the chord within the scale, since all component tones are drawn from the diatonic scale, which contains different intervals between scale steps. Chord type also depends on whether the key is major or minor, giving the distinctive qualities of the major and minor modes. Often a fourth tone is also sounded which repeats one of the three triad components in another octave.

Music theorists have also described a hierarchy of stability for the basic set of chords. The tonic (I) chord is the most stable, followed by the chords built on the fifth (V), fourth (IV), sixth (VI) and second (II) scale tones. The chords built on the third (III) and seventh (VII) scale tones, although within the basic set of chords from a key, are less essential for establishing the key, and tend to appear less frequently and with less rhythmic or metrical stress. In addition chords outside the prevailing key (called non-diatonic) are sometimes employed, but these produce very unstable effects, demanding resolution to the more stable harmonies of the system. The use
of chords, then, is highly constrained by membership in the key, and the sequencing of chords tends to follow very regular patterns. Particular chord sequences, called cadences, are routinely used as formulae for establishing keys and signaling phrase endings in music.

The empirical investigation of harmonic structure

The relative success of comprehensive theories of harmonic structure and the fact that they quantify over abstract entities (chord functions) suggest that the internal representation of harmony is a fruitful focus of study. We argue that the experimental study of this structure as perceived by listeners familiar with the musical tradition is a necessary step toward characterizing the process of music perception. Music theorists and researchers in non-experimental branches of cognitive science consult their intuitions in developing theories about the psychologically salient characteristics of musical organization. However, in the absence of careful empirical investigation we cannot be certain that the average listener is extracting all or any part of this structure. This point is even more crucial for music than it is for language for the following reason. Everybody produces novel sentences in a language. Since linguistic competence is a necessary condition for the production of novel sentences, we can infer such competence in every speaker. We cannot infer musical competence in the majority of people, who do not produce novel musical sequences in composition or improvisation. Most people do listen to music, however, and there is strong reason to suppose that they 'comprehend' it, that is, extract some unifying underlying structure. In support of this is the fact that the acceptance of unstructured or atonally structured music has been limited, even though various composers have attempted to introduce music based on other than tonal structures in the present century. It would seem, then, that most listeners do have considerable knowledge of structure in music.

The present investigation focuses on the knowledge that listeners familiar with Western music have about the harmonic functions of chords. In the first study reported here, half the listeners had some formal training in music theory, in addition to training on an instrument or voice. These listeners had explicit knowledge of music theory. The remaining listeners had no music theory background, although all had instruction on an instrument or voice. Listeners in this second group are comparable to native speakers who have implicit but not explicit knowledge of the language. The pattern of responses of these two groups of listeners provides an
indication of the degree of correspondence between the intuitions of listeners with and without explicit knowledge of music theory. The second study employed almost exclusively listeners with musical experience but with no music theory training. The choice of these subject populations was based on a concern for assessing the effect of explicit knowledge of music theory in the first study, and for investigating rather precise characteristics of implicit knowledge about harmonic structure in both studies. Beyond the scope of the present paper is the question of whether the harmonic principles identified here would be evident for listeners with less or even no experience with the Western musical tradition.

The experimental study of harmonic structure has been difficult in the past because of confounding with melodic factors. This is because the outermost voices in a sequence of chords tend to have salient melodic properties. For instance, the topmost tones typically outline a melody or theme. An additional problem with studying chords is the existence of inversions, which are variants of a basic chord function that depend on which component tone of the chord is in the bass; the root tone of a chord need not be in the bass. In order to minimize these effects, Krumhansl et al. (1982) used chords with component tones sounded over a five octave range. The amplitudes of the components were determined by an amplitude envelope, such that the loudness gradually increased over the first octave and a half, was constant over the center two octaves, and decreased symmetrically over the highest octave and a half. Thus, at both the high and low ends of the frequency range, the loudness tapered off to threshold (after Shepard, 1964), producing chords without clearly defined lowest and highest pitches. The results of that study will be described in detail later. This method, which minimizes melodic factors and obscures inversions, will also be used in the present study.

Principles governing relationships between chords with and without tonal contexts

In this section we outline a number of principles that we presume to govern the perceived relationships between chords. Two basic kinds of principles, formulated in terms of psychological distance, will be stated: those that govern relationships between chords either in no context or in a random context (context independent principles), and those that depend on the establishment of a tonal context (context dependent principles). The context independent principles are strongly suggested by data from Krumhansl et al. (1982), and also embody some very basic ideas in music theory.
In a multidimensional scaling study of the perceived relatedness between chords from three closely related keys, Krumhansl et al. (1982) found that 1) chords from a given key cluster together, 2) the I, IV and V chords form a cluster within the group of chords from each key, and 3) a chord pair in which the second chord is either I, IV or V and the first chord is not is judged more closely related than the same pair in the opposite order. In the present paper we have formulated these three principles as Key Membership, Intrakey Distance and Intrakey Asymmetry, respectively. Although none of these principles is stated in quite this manner in music theory, they nevertheless describe very fundamental music theoretic concepts.

Our purpose is not, however, to simply formalize music theoretic principles. Principles of music theory describe the structure of musical compositions and reflect the intuitions of composers and theorists. We wish to formulate and test psychological principles governing the perception of music by listeners with substantial exposure to music but with little explicit knowledge of musical structure. Toward this end, music theory can provide hypotheses, since it is hoped that psychological and music theoretic principles would show considerable convergence.

The context dependent principles are proposed primarily on the basis of data concerning the perception of tones rather than chords. The perceived organization of tones is largely contextual. Krumhansl (1979) found that the tones that form the tonic triad of the tonal context are perceived as most closely related to each other and the nondiatonic tones least closely related, and that a pair of tones is perceived as more closely related when the first tone is nondiatomic and the second diatomic than vice versa. We hypothesize similar principles for the domain of chords: Contextual Distance and Contextual Asymmetry. Finally, the principle of Contextual Identity below is based mainly on the finding that a tonal context increases the recognizability of a diatomic tone in that context (Krumhansl, 1979).

We will refer to the chords I–VII as the chords that are members of a key, chords that function in a key, or simply as chords from a key. Strictly speaking, all chords have a harmonic function in any given key. However, we will simply deal with those triads built upon the seven notes of the diatomic scale, with all the triad tones drawn from the scale. We will refer to the prevailing context as either tonal or random. In a tonal context, a given key has been established as the prevailing tonality. Chords from this key will be referred to as being in the context key, chords not from this key as being out of the context key. In a random context, or with no context at all, no single key has been established as the prevailing tonality. The set consisting of the I, IV and V chords will be referred to as the harmonic core.
In order to precisely state the harmonic principles, let the psychological distance between two chords, \( C_1 \) and \( C_2 \), in temporal succession be denoted by \( d(C_1, C_2) \). Let \( K \) denote the basic set of seven chords that are members of a key, and let \( C \in K \) mean that the chord \( C \) is a member of \( K \), and \( C \notin K \) mean that it is not from \( K \). The set, \( K \), contains the seven triads built upon the seven degrees of the diatonic scale; the set, \( S \), contains the three chords (I, IV, and V) of the harmonic core.

There are three principles governing context independent relationships. The first principle, Key Membership, states that chords from the same key are perceived as more closely related than chords that are not from the same key, written as:

\[
(1) \quad (\text{Key Membership}) \quad d(C_1, C_2) < d(C_3, C_4), \quad \text{where} \quad C_1, C_2 \in K, \quad \text{and} \quad \text{there does not exist any} \quad K' \quad \text{such that} \quad C_3, C_4 \in K'.
\]

For instance, this principle would assert that the D and E major chords will be perceived to be more closely related (less distant) than the B and D major chords. The D and E major chords are members of the same key (they are, respectively, the IV and V chords in the key of A major), whereas there is no key of which both the B and D major chords are members. These distance relations obtain even when no key has been established as the prevailing tonal context.

The second and third principles characterize the relations between the more stable chords in the key (those in the harmonic core) and the less stable chords of the same key. The second principle, Intrakey Distance, states that chords in the harmonic core, \( S \), are perceived as more closely related to each other than are the other chords from the key but not in the core. This is written as:

\[
(2) \quad (\text{Intrakey Distance}) \quad d(C_1, C_2) < d(C_3, C_4), \quad \text{where} \quad C_1, C_2 \in S, \quad C_3, C_4 \notin S, \quad \text{and} \quad C_1, C_2, C_3, C_4 \in K, \quad C_3 \neq C_4.
\]

The third principle, Intrakey Asymmetry, states that two chords from the same key are perceived as more closely related if the first chord is not in the harmonic core and the second chord is in the harmonic core than if they are heard in the reverse temporal order, or:

\[
(3) \quad (\text{Intrakey Asymmetry}) \quad d(C_1, C_2) < d(C_2, C_1), \quad \text{where} \quad C_1 \in S, \quad C_2 \in S, \quad \text{and} \quad C_1, C_2 \in K.
\]

Consider the key of C major. The C, F and G major chords constitute the stable harmonic core. According to the principle of Intrakey Distance, the F and G major chords (which are, respectively, the IV and V chords in the key of C major) would, for example, be perceived as more closely related
The representation of harmonic structure in music

than, say, the D and E minor chords (which are, respectively, the II and III chords in the same key). The principle of Intrakey Asymmetry would predict, for example, that when the G major chord (which is in the harmonic core) follows the D minor chord (which is not in the core), the two are perceived as more closely related than when they occur in the reverse temporal order.

The first three principles are presumed to hold even in the absence of an established tonal context, i.e., even when no specific key has been set up as the most strongly suggested. The next three principles describe relationships that are altered by a tonal context. Whereas the first three principles characterize the relative stability of chords from the same key, the next three principles characterize the relative stability of chords in a way that depends on their relation to the context key. Let $d_K$ refer to the psychological distance between chords when the context instantiates a key, $K$. Thus, if $C$ is a member of $K$ and $K$ is the key of the context, $C$ will be said to be in the context key (i.e., $C$ is a diatonic chord). If $C$ is not a member of the key of the context, $C$ will be referred to as out of the context key (i.e., $C$ is a nondiatomic chord).

The fourth principle, Contextual Identity, states that if a chord is in the context key, it will be perceived as more closely related to itself (less distant from itself) than if there is no tonal context or if it is out of the context key. Formally, this is:

$$\text{(Contextual Identity)} \quad d_K(C_1, C_1) < d(C_1, C_1) \quad \text{and} \quad d_K(C_1, C_1) < d_K'(C_1, C_1), \quad \text{where} \quad C_1 \in K \quad \text{and} \quad C_1 \notin K'.$$

The perceived relatedness between a chord and itself is equivalent to its ease of recognition. A chord (say, G major) is more likely to be recognized as having occurred earlier if it is in the key of the context (say, the key of C major) than if either there is no tonal context or it is out of the key of the context (say, the key of F major).

The fifth principle, Contextual Distance, governs the perceived relationships between different chords. It states that two chords will be perceived as most closely related if they are in the context key, moderately related when no tonal context is provided, and more distantly related if neither chord is in the context key. This may be written as:

$$\text{(Contextual Distance)} \quad d_K(C_1, C_2) < d(C_1, C_2) < d_K'(C_1, C_2), \quad \text{where} \quad C_1, C_2 \in K, \quad C_1, C_2 \notin K'.$$

As an example, the D and E major chords are perceived as more closely related if embedded within a context of A major (a key of which they are both members) than if heard with no tonal context at all. Furthermore,
they are perceived as more closely related with no tonal context than in some other context key, say C major, a key of which neither is a member.

The sixth principle, Contextual Asymmetry, applies to chords from different keys. It states that a pair of chords will be judged as more closely related if the first chord is out of the context key and the second chord is in the context key than if they are heard in the reverse temporal order. In addition, an earlier sounded chord outside the key will be confused more frequently with a later sounded chord from the key than the opposite temporal order. The Contextual Asymmetry principle is written as:

\[(6) \text{(Contextual Asymmetry)} \quad d_K(C_1, C_2) < d_K(C_2, C_1), \quad \text{where} \quad C_1 \notin K \quad \text{and} \quad C_2 \in K.\]

If the context key is A major, for instance, a C major chord (nondiabatic in A major) will be less stable than a D major chord (diatonic in A major), and thus the C major chord will be judged as more closely related to and confused more frequently with a D major chord than the opposite.

Having stated these six principles, we turn now to two experiments designed to test them using two quite different measures of psychological distance. The first was a scaling study of the chords of two distantly related keys, in which listeners judged chord pairs presented either with no context or with a tonal context. The second experiment investigated confusions between chords embedded in tonal or random sequences in a recognition memory paradigm. We elected to use these two quite different methods in order to evaluate the generality of the principles just stated.

**Experiment 1**

In an earlier study of the perceived relationships between musical chords, Krumhansl et al. (1982) employed the chords from three related keys: C major, G major (which is the key built on the dominant, V, of C major), and A minor (which is the relative minor of C major). That study was concerned with determining the nature of the psychological representation of chords, particularly as it supports the assimilation of modulations (changes) between closely related keys. Listeners heard all possible ordered pairs of chords from these keys following an ascending scale in one of the three keys that varied between blocks of trials. Listeners rated how well the second of the two chords followed the first, in the context of the ascending scale. Multidimensional scaling of the resulting judgments produced a compact central cluster of the I, IV, and V chords from the three keys. In addition, the chords that are shared by at least two of the
The representation of harmonic structure in music

The seven chords from each key were found to have a hierarchical structure. The hierarchical clustering method applied to the data grouped the chords from each key together in the following order: I, V, IV, VI, II, III, and VII. This ordering reflected the hierarchy of stabilities within the basic set of harmonies described by music theory. In addition, asymmetries were found, such that pairs ending on a chord from C major were preferred to pairs ending on a chord not from C major. However, the perceived relationships between chords were not for the most part found to depend on the key of the context scale immediately preceding the chords. This may have been a consequence of the fact that the three keys were selected to be closely related; either the experimental measure was insufficiently sensitive to reflect these subtle key differences, or, owing to the considerable harmonic overlap of both G major and A minor with C major, the listeners heard the key of C major as the prevailing tonality throughout the experiment. Alternatively, the scale context may have been insufficiently strong to instantiate the different keys; the test chords, as musical entities with considerable tonal influence in their own right, may have overridden the contextual effects of the scales.

These results contrast sharply with the results of an earlier scaling study of single tones presented in a tonal context (Krumhansl, 1979). In that study, pairs of tones were presented following a scale context, and the judgments reflected a great deal of structure that was dependent on the tonality of the scale context. Multidimensional scaling of the ratings of all possible ordered pairs of tones yielded a conical structure in which the three tones of the tonic triad of the context were nearest the vertex, the other diatonic tones were farther from the vertex (and thus farther from each other), and the nondiatonic tones were neither closely related to the diatonic tones nor to each other. The judgments contained a very regular pattern of asymmetries, such that pairs ending on the most stable tones of the instantiated key were preferred to pairs ending on less stable tones. Thus, very strong context-dependent effects were found for single tones.

In order to explore whether context also affects the perceived relationships between chords, the present experiment used chords from two major keys, C major and F# major, that music theorists would classify as maximally unrelated. These keys have the greatest difference in terms of their
key signature, with C major having no sharps or flats and F# major having six sharps. In addition, they share only one scale note, and have no chords in common. Thus, there are a total of fourteen different chords from the two keys, seven from each. The chords from C major are: C major (I), D minor (II), E minor (III), F major (IV), G major (V), A minor (VI), and B diminished (VII). The chords from F# major are: F# major (I), G# minor (II), A# minor (III), B major (IV), C# major (V), D# minor (VI), and E# diminished (VII). The present experiment employed all possible ordered pairs of these chords, which were sounded either without any context or in the context of one of the two keys. A tonal context was established by preceding the chord pair by a IV V I cadence in the key of either C major or F# major. Cadences rather than scales were used because they may be stronger instantiators of the intended context keys. In this study, we focus on both context-dependent and independent relationships that are perceived between the chords of maximally distant major keys.

Method

Subjects

Sixteen subjects (eight male and eight female), all from the Cornell summer community, participated for an hour and a half over two sessions. All subjects except the author, J. B., were paid $5.00 for their participation. On the average, subjects had played musical instruments or sung for 11.4 years, and currently played or sang for 4.8 hours per week. Half the subjects had never taken a course in music theory, and half had studied some music theory at the introductory or intermediate levels. All subjects reported having normal hearing, and none reported absolute pitch.

Apparatus

The stimuli were generated by a Hewlett Packard 1000L computer, programmed to produce the chords by creating digital representations and playing them out through a Hewlett Packard (59303A) digital-to-analog converter. An A.P. Circuit Corporation Variable Frequency Filter (Model AP-255-5) was used to eliminate high frequency noise introduced by the conversion process. Recordings were made on Maxell UD 35-90 tapes on a Revox A77 tape recorder. Subjects listened to the tapes on the same tape recorder at 7.5 inches per sec (19 cm per sec) through an Ampex AA-620 loudspeaker at approximately 67 dBA sound pressure level.
Stimuli

The tones comprising the chords were all drawn from the set of equally spaced semitones on a logarithmic scale with twelve semitones per octave, as would be obtained by an equal tempered tuning based on 440 Hz (A). An amplitude envelope was imposed on the chord components so that the tones at the low and high ends of the five-octave range tapered off to threshold. Details of this method are described in Krumhansl et al. (1982). Each chord lasted 500 msec, with 10 msec amplitude rise and fall times to reduce onset and offset clicks. Approximately 250 msec of silence separated the three successive cadence chords from each other and the two test chords from each other. There was a 750 msec pause between the end of the context cadence and the beginning of the test chords, and a 4 sec interval between trials. After every ten trials, a 500 msec burst of white noise signaled the beginning of another group of ten trials.

Procedure and experimental design

In the two tonal context conditions, each trial consisted of a three-chord cadence (IV V I) in the context key (C or F# major), followed by a pair of test chords in succession. In the no context condition, the pairs of chords were presented in isolation. There were three blocks of trials, one for each context condition. Subjects first heard the two tonal context conditions in random order, followed by the no context condition. Each block contained 182 randomly ordered trials, exhausting all possible ordered pairs of non-repeating chords selected from the set of fourteen chords (seven from the key of C major and seven from F# major). Each block was preceded by ten practice trials, which were the same as the last ten trials of the block. Subjects were instructed to rate on a seven point scale how well the second chord followed the first. In the two tonal context conditions, these ratings were to be made with respect to the context of the three-chord cadence at the beginning of each trial.

Results

Individual differences

The relatedness judgments of all ordered pairs of chords averaged across contexts were represented in a 14 × 14 matrix minus the diagonal entries. Intersubject correlations were computed with these matrices. The average
intersubject correlation was 0.438. All intersubject correlations were significant \((p < 0.05)\), except three intersubject correlations all of which involved the same subject. However, that subject’s responses correlated significantly with the remaining 12 subjects. The average intersubject correlation among the eight subjects with music theory was 0.462, and among the eight without music theory was 0.407. Neither of these was significantly different from the average intergroup correlation, which was 0.441. Therefore, based on these correlations it does not appear that there were substantial systematic differences between subjects with and without formal instruction in music theory. The multidimensional scaling and hierarchical clustering methods were applied to the matrix obtained by averaging across all subjects.

**Multidimensional scaling**

For each context condition, the \(14 \times 14\) matrix of relatedness judgments was averaged across the diagonal before application of the multidimensional scaling method (Kruskal, 1964a,b; Shepard, 1962). The particular program used was MDSCAL (Kruskal, Reference Note 1). Figure 1 shows the two-dimensional solutions for each of the three context conditions. The result for the no context condition, shown at the top of Figure 1, had stress value (Formula 1) = 0.156. In this configuration, the chords were perfectly separated into two groups corresponding to the two keys from which they were drawn. In the solution for the C major context, shown on the left of Figure 1 (stress = 0.158), the chords from C major were pulled closer together and those from F# major were more dispersed than in the no context condition. Similarly, in an F# major context, shown on the right of Figure 1 (stress = 0.195), the chords from F# major were pulled together and those from C major were more separated. This illustrates the Contextual Distance principle. Although the stress values were rather high, the configurations obtained in the higher-dimensional solutions did not contain additional interpretable dimensions or clusters.

In all three frames of Figure 1, the fourteen chords separated into two clusters representing the two keys. Thus, regardless of context, chords from a single key were judged to be more closely related to each other than chords not in the same key, confirming the Key Membership principle. This effect is independent of context.

In addition to the global separation of chords according to key membership, the solutions exhibited a fairly stable structure within the set of chords of each of the two keys. Generally, the I, IV, and V chords occupied central positions within each key. This Intrakey Distance effect was most evident
Figure 1. Multidimensional scaling solutions of fourteen chords—seven from the key of C major and seven from the key of F# major. All ordered pairs of chords were judged (a) with no context (top), (b) in a C major context (left), and (c) in an F# major context (right). Chords separate according to key membership. A tonal context shrinks distances between chords in the context key, and stretches distances between chords out of the context key.

for the chords from C major in either a C major context or no context, and the chords from F# major in either an F# major context or no context. The effect appeared less pronounced for the chords from C major in an F# major context or the chords from F# major in a C major context. Multidimensional scaling of the seven chords from each key separately produced a cluster containing the I, IV and V chords in the center of the two-dimensional solutions, surrounded by the II, III, VI and VII chords. This pattern was found when the chords were in the context key, when there was no context, and also when the chords were out of the context key. This shows that the Intrakey Distance principle is independent of context.

The half-matrices used for multidimensional scaling of the fourteen chords were intercorrelated to supplement visual inspection of the solutions.
The correlations of the no context matrix with matrices of the C major and F# major contexts were 0.614 and 0.566, respectively, and the correlation of the C and F# major context matrices was -0.150. Thus, large differences were found between the matrices of the C and F# major contexts, with the no context matrix more similar to both. In addition, the C and F# major context matrices can be made to match each other with respect to the chords that are diatonic or nondiatonic. For instance, chords from C major in the context key of C major may be considered equivalent to chords from F# major in the context key of F# major. Since music theory would predict that C and F# major are maximally distant keys (they have positions on the circle of fifths that are diametrically opposite each other), chords from C major in the context key of F# major may be considered equivalent to chords from F# major in the context key of C major. Thus, one matrix can be restructured so as to match the in- versus out-of-context-key structure of the other. The correlation of the C major context matrix with the restructured F# major context matrix was 0.940. This demonstrated that the C major context had the same effect on the chords of C and F# major as did the F# major context on the chords of F# and C major, respectively.

Hierarchical clustering

Hierarchical clustering (Johnson, 1967) was used to further examine the structure of intrakey relationships. This technique complements multidimensional scaling, and showed that there was considerable consistency in the fine structure of chord relationships within the harmonic core (I, IV and V) and outside the core (II, III, VI and VII). Figure 2 displays the clustering solution (using the compactness method) for the average of the four seven × seven matrices representing the relationships between the chords when the chords were in the context key and when there was no context (i.e. chords from C major in a C major context and no context, and chords from F# major in an F# major context and no context). The I and V chords clustered first, followed by the IV chord. This cluster, the harmonic core, was then joined by the VI and II chords, and finally by the VII and III chords. Although clustering of the chords outside the context key also showed the I, IV, and V grouping together before the others, the specific hierarchies within and outside the harmonic core were much less consistent than when the chords were in the context key or when there was no context.
The representation of harmonic structure in music

Figure 2. Hierarchical clustering of the seven chords (I-VII) from a key (when the chords were in the context key and when there was no context). The I and V chords cluster first, followed by the IV chord. The VII and III chords cluster last.

**Contextual and Intrakey Distance**

The left panel of Figure 3 shows the average ratings for chord pairs from C major and from F# major in each of the three context conditions. Ratings were highest when the pair of chords was in the context key, intermediate when there was no context, and lowest when the pair was not in the context key. A 3 x 2 repeated measures analysis of variance yielded a significant main effect of context ($F(2, 15) = 12.64, p < 0.001$) and a significant context x key membership interaction ($F(2, 30) = 152.13, p < 0.001$). This interaction demonstrates the extent to which the perceived relatedness of chords from a key is dependent upon context. Chords from a key were judged to be more closely related to each other when there was a tonal context that established the same key than when there was no tonal context; and chords from a key were judged to be less closely related to each other when there was a tonal context that established an unrelated key than when there was no tonal context. The inequalities of the Contextual Distance principle were therefore found to be significant.

In contrast, the right panel of Figure 3 shows that the relative difference between the perceived relatedness of chords in and out of the harmonic core of a key were much less affected by context. The I, IV and V chords were judged to be more closely related to each other than were the ii, iii, VI and VII chords. This was true when the chords were in the context key, when there was no context, and even when they were out of the
Figure 3. **Contextual and Intrakey Distance.** Left: Perceived relatedness of chords from the key of C major and from the key of F# major, with either a C major context, no context, or an F# major context. Chords in the context key are judged more closely related to each other than chords out of the context key. Right: Perceived relatedness of chords within the harmonic core (I, IV, V) and chords outside the harmonic core (II, III, VI, VII) when the chords are in the context key, with no context key, and when the chords are out of the context key. In all context conditions, chords within the harmonic core are judged more closely related to each other than chords outside the core.

context key. A repeated measures analysis of variance of the three context conditions (in context, no context, out of context) and whether or not the chord pair was a member of the harmonic core (I, IV, V) yielded a significant main effect of membership in the core ($F(1, 15) = 91.16, p < 0.001$), confirming the Intrakey Distance principle. In addition, there was a significant main effect of context condition ($F(2, 38) = 68.35, p < 0.001$), as would be predicted by the Contextual Distance principle. A significant interaction of core membership with context condition ($F(2, 30) = 6.53, p < 0.01$) reflects the finding that the advantage of the harmonic core chords over the others was reduced when the chords were out of the context key. However, even when the chords were out of the context key the ratings for core chords were higher than for chords not in the core, thus supporting the Intrakey Distance principle.

**Contextual and Intrakey Asymmetry**

In order to determine whether there were asymmetries that depended on key membership, the average rating across subjects was computed for the 49 pairs of chords with one chord in C major and one chord in F# major. A three-way analysis of variance was computed for these scores with the following three factors: context (C major, F# major, or no context), temporal order (whether the first chord was in C or F# major), and chord pair
Figure 4. **Contextual and Intrakey Asymmetry.** Left: Asymmetries in the perceived relatedness of chord pairs in which one chord was from the key of C major and the other from the key of F# major. A chord in the context key is perceived as following a chord out of the context key better than the reverse temporal order. Right: Asymmetries in the perceived relatedness of chord pairs in which one chord was from the harmonic core and the other chord was not. In all context conditions, a chord from the harmonic core is perceived as following a chord from outside the harmonic core better than the reverse temporal order.

(which constituted the repeated measures factor of the design). The results are shown on the left of Figure 4 as a function of context. There was a significant main effect of context ($F(2, 96) = 26.54, p < 0.001$), such that the highest ratings were given when there was no context. Thus, the tonal context tended to increase the distance between chords that belong to different keys. There was a strong interaction between context and temporal order ($F(2, 96) = 83.74, p < 0.001$); higher ratings were given to pairs that ended on a chord in C major when the context was C major, and just the opposite when the context was F# major. In other words, higher ratings were given to chord pairs that ended with a chord in the context key. No differences between the two orders were found when there was no context. This interaction confirms the Contextual Asymmetry principle.

In contrast, asymmetries in chord relationships within the same key were much less influenced by context, as shown on the right of Figure 4. Chord pairs in which the first chord was not in the harmonic core (i.e., II, III, VI or VII) and the second chord was in the harmonic core (i.e., I, IV, or V) were judged to be more closely related than pairs in the reverse temporal order ($F(1, 11) = 17.93, p < 0.005$). Both temporal orders were influenced in the same way by context, as seen by a main effect of context ($F(2, 22) = 165.20, p < 0.001$). There was a significant interaction of context condition with temporal order ($F(2, 22) = 6.07, p < 0.01$), indicating that the asymmetries were reduced in magnitude when the chords were out of the
context key. The functions do not cross, however, generally supporting the Intrakey Asymmetry principle.

Discussion

In this experiment, the relatedness judgments of chords contained very regular patterns that depended on key membership and the functions of chords within keys. Moreover, strong contextual effects were found in these judgments. The results showed that chords, like single tones, are subject to influences of the tonal context in which they are embedded. The results may be summarized with reference to five of the principles stated earlier for harmonic relationships.

The first principle tested was Key Membership, which states that even when no context is provided, the psychological distance between chords from a single key is less than that between chords from different keys. This was seen in the multidimensional scaling solution as the clear separation of chords from the keys of C major and F# major into two quite distinct clusters. These results extend the earlier study of Krumhansl et al. (1982), which showed that the chords from different though closely associated keys tended to be perceived as more distantly related than chords from a single key. This separation was found in the present experiment to be even more marked when the chords were drawn from distant keys. Thus, influences from the level of key relationships are clearly evident at the level of chords: the internal representation of chords reflects the relative proximity of chords from a single key or from closely related keys as compared to chords from distantly related keys.

The present results showed a hierarchy within each key similar to the one obtained by Krumhansl et al. (1982); the I, IV, and V chords of each key (the members of the harmonic core) were grouped more closely than the remaining diatonic chords, as predicted by the Intrakey Distance principle. This was found even in the absence of a tonal context, and confirms the hierarchy of harmonic functions described by music theorists. The listeners have apparently internalized the prominence given to the most harmonically stable chords—those in the harmonic core—in tonally structured musical sequences.

Chords in the harmonic core were found to have another property that marks their stability: chord pairs in which the first chord was not in the harmonic core and the second one was were preferred to pairs heard in the reverse temporal order. This is described by the Intrakey Asymmetry principle, which was found to be largely independent of context. Asym-
metric relations have been explained by postulating the existence of certain prototypical elements as cognitive reference points (Rosch, 1975). We suggest that the chords in the harmonic core function as cognitive reference points, as will be discussed later. The Intrakey Asymmetry principle was found to be less robust, though still operative, when the chords were outside the context key.

The Contextual Distance principle was tested by comparing the average rating for pairs of chords from a single key as a function of the context condition. When the two chords were drawn from the key of the context, they were perceived as more closely related than when there was no context, and the lowest ratings were given when the chords were not drawn from the context key. In the multidimensional scaling solution, chords in the context key pulled together and chords out of the context key separated from each other. From this we can conclude that chords, like single tones, are perceived according to their relationships to the prevailing tonality; the failure to obtain context effects in the earlier experiment (Krumhansl et al., 1982) most likely resulted from the selection of chords from three closely related keys.

The Contextual Asymmetry principle was strongly confirmed by comparing the ratings given to pairs of chords from two different keys as a function of context. When the first chord was outside the context key and the second chord within it, higher ratings were given than when the same two chords were heard in the opposite temporal order. When no context was provided, no asymmetries depending on key membership were found. Thus, the cadence context serves to establish the chords of the key as stable reference points (Rosch, 1975) toward which nondiatonic chords are drawn. In music theory this is described as the resolution that is demanded by nondiatonic harmonies.

The present results strongly confirmed five of the six principles that were presumed to describe the internal representation of harmonic information. The remaining principle, Contextual Identity, could not be tested because the present experiment did not include pairs with repeated chords. This principle was tested in the second experiment, which used recognition memory performance as the dependent measure.

The primary motivation for the second experiment, however, was to test the hypothesis that the representational assumptions do in fact have strong implications for music perception, and are directly rooted in psychological processes. In this first experiment, listeners were instructed to rate how well the second chord followed the first, and to make these judgments in the tonal context when one was provided. These instructions may well be criticized because of their open-ended nature: exactly what aspects of the
harmonic structure are to be considered, and how is the key of the context to be taken into account? Despite this vagueness, reasonable consistency across listeners was found. Moreover, since no regular differences appeared between the responses of listeners with and without backgrounds in music theory, it could be concluded that the results did not simply depend on explicit training in harmonic structure. The results therefore seem to reflect the implicit knowledge listeners with experience with the Western musical tradition have about harmonic information.

Nonetheless, given the artificial nature of the task, converging evidence from a rather different task is desirable. We chose to look at recognition memory for chord sequences. Memory is clearly necessary for the apprehension of musical structure. For example, the sense of key which develops and is maintained over a sequence of musical events requires memory for past musical events. The apprehension of larger musical forms necessitates the ability to recognize themes or motifs. If systematic memory effects are found in the experimental situation, they would presumably reflect properties of the representation of harmonic information. Our characterization of these properties would be strengthened if we could adduce similar properties from two very different kinds of data: relatedness judgments and recognition memory judgments.

Experiment 2

Recognition memory for tones has been found to depend on the relationship of the to-be-remembered tone or tones to the other tones that appear in the context. Specific interfering effects of extraneous or masking tones in the same pitch range have been firmly established (Deutsch, 1972a, 1973a, 1974; Divenyi and Hirsh, 1975; Idson and Massaro, 1976). Deutsch (1972b) and Divenyi and Hirsh (1978) found that recognition memory for tone sequences is further impaired when the intervals of the sequence were expanded, and Deutsch (1978) reported that recognition accuracy was inversely related to the average size of the intervals contained in a sequence of interfering tones. In these cases, the masking tones were not selected for their musical relationships to the to-be-remembered tones.

The importance of musical structure on memory for tones has been demonstrated in a number of studies. For example, Deutsch (1973b) showed that specific memory interference effects generalized across octaves. Moreover, Divenyi and Hirsh (1974) found better identification of temporal order when the frequencies formed musical intervals such as major thirds,
minor sixths, and perfect fourths and fifths. One experiment in an important collection of studies by Frances (1972) showed that tonal melodies (conforming to diatonic intervals) were better remembered than were random melodies, a result replicated by Dewar, Cuddy and Mewhort (1977) and Cuddy, Cohen and Miller (1979). Moreover, this last study showed that changes in a three-tone sequence that were outside a diatonic scale were easier to detect when the three tones were embedded in a tonal context than when presented in isolation. This last result was interpreted as indicating that listeners use rules defining significant musical relationships to detect structural violations. A somewhat similar result was obtained by Dowling (1978), who showed that changed intervals in tonal answers (that conformed to the diatonic scale of the key of the to-be-remembered sequence) were very difficult to detect. These results may be summarized as follows. In general, tonal sequences are easier to remember, as evidenced by better recognition of exact repetitions or transpositions and easier detection of randomly changed tones, than are less well-structured musical sequences. However, changes that conform to the tonal structure are difficult to detect.

The present study employs sequences of chords, rather than single tones, to determine the generality of these memory effects and to provide a further test of the principles of harmonic structure stated earlier. The listeners first heard a sequence of seven chords (called the standard sequence) followed by a second sequence of seven chords (called the comparison sequence). The two sequences were either identical, chord for chord, or one chord (called the target chord) was changed between the two sequences. The listeners were required to make a ‘same’ or ‘different’ response for each standard-comparison pair.

The sequences were either random, with the chords drawn randomly from different keys, or tonal, with all chords drawn from a single major key (except possibly the target chord which could be nondiatonic). The selection of chords for the tonal sequences was restricted to the set of chords with relatively stable harmonic functions in the key: I, II, IV, V and VI. Each sequence had two repeating chords, which were never in successive positions in the sequence. Moreover, the tonal sequences were constructed using the rules of harmonic progression described by Piston (1962, p. 18), and the last chord in each tonal sequence was the tonic (I) chord of the key. The rules of harmonic progression required that the penultimate chord be V or IV, so each tonal sequence ended with a final V I or IV I cadence. The random sequences were constructed to match the tonal sequences in terms of the number of repetitions and the number of major and minor chords.
Table 1. **Conditions of Experiment 2**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Type</th>
<th>Context</th>
<th>Target chord in standard sequence</th>
<th>Target chord in comparison sequence</th>
<th>Percent correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Different</td>
<td>Tonal</td>
<td>Diatonic</td>
<td>Diatonic</td>
<td>0.569</td>
</tr>
<tr>
<td>2</td>
<td>Different</td>
<td>Tonal</td>
<td>Diatonic</td>
<td>Nondiatic</td>
<td>0.868</td>
</tr>
<tr>
<td>3</td>
<td>Different</td>
<td>Tonal</td>
<td>Nondiatic</td>
<td>Diatonic</td>
<td>0.613</td>
</tr>
<tr>
<td>4</td>
<td>Different</td>
<td>Tonal</td>
<td>Nondiatic</td>
<td>Nondiatic</td>
<td>0.776</td>
</tr>
<tr>
<td>5</td>
<td>Different</td>
<td>Random</td>
<td>Random</td>
<td>Random</td>
<td>0.667</td>
</tr>
<tr>
<td>6</td>
<td>Same</td>
<td>Tonal</td>
<td>All chords diatonic</td>
<td></td>
<td>0.693</td>
</tr>
<tr>
<td>7</td>
<td>Same</td>
<td>Tonal</td>
<td>One chord nondiatic</td>
<td></td>
<td>0.550</td>
</tr>
<tr>
<td>8</td>
<td>Same</td>
<td>Random</td>
<td>All chords random</td>
<td></td>
<td>0.550</td>
</tr>
</tbody>
</table>

The eight conditions of the experiment are shown in Table 1. In the first five conditions, the standard and comparison sequences differed in terms of one chord, the target, which could appear in serial positions two through six. In condition one, a diatonic chord in the first sequence was replaced by another diatonic chord in the second sequence in the same serial position. This change was made in such a manner as to preserve the number of repetitions and ensure that the resulting sequence again conformed to the rules of harmonic progression. In the second condition, a diatonic chord in the first sequence was changed to a nondiatomic chord in the second sequence. In the third condition, the target chord was changed from a nondiatomic to a diatonic chord, and in the fourth condition, the nondiatomic target chord was changed to another nondiatomic chord. The fifth condition used random sequences in which one chord was changed to a randomly selected chord so as to match the pattern of repetitions and chord types of the first four conditions. Conditions six, seven and eight were exact repetitions. Identical diatonic chords appeared in the corresponding positions of the standard and comparison sequences in the sixth condition, and a single nondiatomic chord was repeated in the standard and comparison sequences (which were otherwise tonal) in the seventh condition. The last condition had identical standard and comparison sequences that were constructed from randomly selected chords.

This design enabled us to study recognition errors as a function of context. Systematic differences would be expected depending upon whether the context is tonal or random, and if tonal, whether the target chords are diatonic (i.e., from the context key) or nondiatomic.
Method

Subjects

Eleven subjects, drawn mostly from Harvard summer school, were paid $14 each for participating in three experimental sessions lasting a total of approximately four hours. All subjects had at least three years instruction on a musical instrument or voice, averaging 8.7 years. Only one subject had any formal training in music theory. Normal hearing was reported by all, and absolute pitch by none.

Apparatus

The stimuli were generated under the control of a PDP-15 computer in the Laboratory of Psychophysics at Harvard University. Digital representations of the chords were played through a digital-to-analog converter and recorded on Maxell UD 35–90 tapes at 7.5 in. per sec on a Sony TC-540 tape recorder. Subjects listened to the tapes through headphones using the same tape recorder.

Stimuli

The method used to construct the chords was the same as for Experiment 1. Each chord sounded for 500 msec with 10 msec amplitude rise and fall times. A 60 msec interval separated chords within a sequence. The interval between the standard and comparison sequences was 560 msec. A 3.5 sec pause between trials allowed the subjects to write their responses. A 600 msec burst of pink noise followed by a 400 msec silence preceded each trial, except that two bursts of pink noise were presented after every group of ten trials to ensure that listeners were at the correct position on the response sheet.

The conditions of the experiment are listed in Table 1. All eight conditions contained 100 trials, except for condition five which contained 200 trials. In the first five ‘different’ conditions the altered chord appeared in the five possible serial positions (two through six) equally often. The 900 trials were randomly intermixed and divided into 10 blocks of 90 randomly ordered trials, subject to the constraint that each serial position of the changed chord in the ‘different’ conditions appear equally often in each block. Each subject heard the blocks in a different random order.
Procedure

Subjects were instructed to judge whether the two sequences were the same or different on a four-point scale (−2 = Very sure same, −1 = Same but not sure, 1 = Different but not sure, 2 = Very sure different). The listeners were told that at most one chord of the sequence would be changed, that it would never be the very first or last chord, and that there were twice as many 'different' as 'same' trials. The first block of trials was preceded by 10 randomly selected practice trials.

Results

The average percent correct for each of the eight conditions was found by dichotomizing the response scale into 'same' and 'different' responses; these values are shown in Table 1. Looking first at the five 'different' trial types, performance was highest in the second condition, in which a diatonic chord was changed to a nondiatomic chord in a tonal context. Thus, changes from chords in the context key to chords out of the context key were most easily detected. This was followed by condition four, in which a nondiatomic chord was changed to another nondiatomic chord in a tonal context. Performance on the random sequences in condition five was next highest. Condition three followed next, in which a nondiatomic chord in the standard sequence was changed to a diatonic chord. The poorest performance of all the 'different' conditions was seen in condition one, in which a diatonic chord was changed to another diatonic chord. From the low performance on conditions three and one, we see that subjects tended to incorrectly judge as 'same' trials on which the changed chord in the comparison sequence was diatonic. Thus, chords in the context key were often incorrectly judged as having previously occurred in the target position in the first sequence.

A two-way analysis of variance of the percent correct scores in the 'different' conditions and five serial positions yielded a significant main effect of condition (F(5, 50 = 17.88, p < 0.001). In that analysis, two different random conditions were distinguished, one which matched the repetition pattern of condition one, and one which matched the repetition pattern of conditions two through four. However, no differences were found between these two trial types, so they were averaged together in Table 1. In addition, a main effect of serial position was found (F(4, 40) = 4.69, p < 0.01), with a tendency for changes later in the sequence to be detected more easily than changes early in the sequence. This effect may depend
The representation of harmonic structure in music

on the fact that the rules of harmonic progression tightly constrain the possible chord substitutions in the penultimate serial position. The interaction between serial position and condition was not significant. Subjects' confidence ratings were used to compute areas under the MOC curves, using the appropriate 'same' conditions as 'noise'. This analysis yielded similar results, except that performance in condition one was somewhat higher than in condition three. No other reversals were found.

A one-way analysis of variance of the three 'same' conditions showed a significant difference between these three conditions ($F(2, 20) = 16.66, p < 0.001$). Performance was highest in condition six, in which all chords were diatonic. There was no difference in the low performances in condition seven, in which one chord was nondiatonic, and condition eight, in which chords were randomly selected. Thus, subjects tended more often to correctly judge as 'same' those sequences in which all chords were diatonic. Since no chords were changed in the 'same' conditions, serial position was not a factor in this analysis.

Discussion

Only the context-dependent principles were explicitly tested in this experiment. Key Membership was not tested since no distinction was made between pairs of target chords from the same key or from different keys in the random sequences. Furthermore, no distinction was made between targets in the harmonic core of a key and those not in the core; hence the other two context-independent principles, Intrakey Distance and Intrakey Asymmetry, were also not assessed.

The more closely related to itself an object is perceived to be (i.e., the less its psychological distance from itself), the more easily it will be recognized. The principle of Contextual Identity would therefore predict that a chord would be more easily correctly recognized when it is in the context key than when there is no tonal context or when it is out of the context key. Significantly higher performance was found in condition six than in the other 'same' conditions, in support of the principle. Membership in the key of the prevailing tonal context seems to increase the stability of the memory representation of a chord. A possible caveat concerning this interpretation is discussed below.

The more closely related two different objects are perceived to be (i.e., the less the psychological distance between them), the more easily they will be confused. The principle of Contextual Distance would thus predict that when two chords are in the context key they will be more easily
confused than when there is no tonal context, and less easily confused when they are out of the context key than when there is no tonal context. Performance was lower (more confusions occurred) in condition one (diatonic targets in a tonal context) than in condition five (random context), which in turn was lower than in condition four (nondiatonic targets in a tonal context). This confirms the principle of Contextual Distance.

According to the principle of Contextual Asymmetry, a chord out of the key of the context will be more easily confused with a chord in the key of the context than vice versa. In other words, in a tonal context, the replacement of a nondiatonic chord by a diatonic chord (condition three) will be more difficult to detect than the replacement of a diatonic chord by a nondiatonic chord (condition two). Performance in condition three was strikingly lower than in condition two, thus confirming the asymmetry. Diatonic chords (chords in the context key) are more stable than nondiatonic chords (chords out of the context key), and the less stable chord is psychologically closer to the more stable chord than vice versa.

It could conceivably be argued that all three principles simply reflect a preference for the second sequence being completely diatonic. In other words, subjects tended to respond ‘same’ when the comparison sequence was completely diatonic, regardless of whether it was the same or different; hence the poor performance in the ‘different’ conditions one and three, and good performance in the ‘same’ condition six. However, the signal detection analysis showed that the differences in performance across conditions reflect significant differences in sensitivity (as measured by the areas under the MOC curves), not simply differences in bias. Thus, the low performance in condition one (as compared with conditions four and five) and the high performance in condition six (as compared with conditions seven and eight) do not simply reflect a bias to say ‘same’ when both sequences are completely diatonic (even though there is also a significant bias). Rather, it is actually more difficult to discriminate between different diatonic chords than between different nondiatonic chords, as stated by Contextual Distance. Likewise, the low performance in condition three (as compared with condition two) reflects a lower sensitivity, in support of Contextual Asymmetry. Signal detection analysis is less unequivocal in the case of Contextual Identity. It is possible to attribute the difference in performance in the ‘same’ conditions to a bias to say ‘same’ when both sequences are completely diatonic; hence the high performance in condition six. Although this explanation cannot be ruled out, it may not necessarily be an alternative that is independent of Contextual Identity, which may be the principle underlying this bias.
An examination of the serial position curves indicated that superior performance for changes made later in the sequence was largely due to high performance when the change was made in the penultimate position. Since only two diatonic chords (IV and V) were permissible in this position, detection of changes in this position may have been easier than in the other positions, in which three or more diatonic options existed.

Each prediction of the proposed principles of harmonic organization that could be tested within the present experimental design was confirmed by the results. Completely tonal sequences were better recognized than random sequences or tonal sequences with one nondiatomic chord (Contextual Identity), though this principle could not be as clearly established as the others. Two chords were most frequently confused with each other when both were diatonic, less frequently confused in a random context, and even less frequently confused when both were nondiatomic chords in a tonal context (Contextual Distance). Finally, a nondiatomic chord was more frequently confused with a diatonic chord than vice versa (Contextual Asymmetry). Thus, very reliable effects on memory performance were obtained which converge on the spatial model derived from the relatedness judgments of the previous experiment.

General Discussion

The internal representation of harmonic relationships was shown in these two experiments to be very regular and highly structured. This representation has both context independent and dependent features. Regardless of context, chords from distantly related keys were separated into distinct subgroups. Chords from closely related keys necessarily form less separated, overlapping subsets because of their multiple harmonic functions (Krumhansl et al., 1982). A central core of the most harmonically stable chords from each key was also found; the I, IV and V chords of each key formed a closely interrelated subset within the set of chords from the key, and the relative stability of these chords was reflected in asymmetries in the relatedness judgments.

When a tonal context was introduced, the representations of chords in the context key were made more stable and chords out of the context key less stable than when no tonal context was provided. This effect had three consequences. First, chords in the context key were more frequently correctly recognized than chords outside that key. Second, the context strengthened the perceived relationships among chords in the context key and weakened those between chords outside the context key, an effect
that was evident in both direct relatedness judgments and memory confusions. Finally, temporal asymmetries were found such that a chord outside the context key was perceived as more closely related to a chord in the context key than the reverse; this effect was also found in confusion errors. Thus, the representation of harmonic information, as reflected in converging evidence from the scaling and memory experiments reported here, contains rich structural information about the harmonic functioning of chords from different keys, even in the absence of a tonal context. This information becomes much more finely differentiated when the chords are embedded within a tonal context.

Considerable work has been done on the effect of structured contexts on the perception of speech. Phonemic confusions depend systematically on the context in which the sounds are embedded. Miller and Isard (1963) found that words accompanied by noise were most intelligible when in the context of correct sentences, less intelligible when in the context of grammatical but meaningless sentences, and least intelligible when embedded in ungrammatical strings. Warren (1970) has demonstrated the perceptual restoration of phonemes replaced by coughing sounds in a sentence context. Mispronunciations of words in the context of prose have been shown by Cole (1973) to be more difficult to detect when the changed phoneme differs by only one distinctive feature (Keyser and Halle, 1968) than when it differs by several distinctive features. In a shadowing task, frequent restoration of grossly mispronounced words was found only for syntactically and semantically well-structured contexts (Marslen-Wilson, 1975). Marslen-Wilson and Welsh (1978) showed that, in a prose context, restoration of mispronounced words by shadowing is more frequent when the changed phoneme differs by fewer distinctive features, when the sentence context is more highly constrained, or when the changed phoneme is in the third rather than first syllable (since the first two syllables further constrain the interpretation of the word).

We refrain, however, from drawing any substantive parallels between phonemic confusions in language and harmonic confusions in music. There are no clear semantic constraints in music, and grammatical constraints take a rather different form (see the section on generative music theory below). Furthermore, there are no concrete analogies between linguistic units (such as words or phonemes) and musical units (such as chords or tones). However, both domains consist of temporal sequences of auditory events, and in each domain our perception of events seems to be systematically influenced by a context with a particular kind of structure. In the present paper we have investigated some of these regularities in the perception of music.
Hierarchies of Stability and Cognitive Reference Points

The perception of coherence in the highly embellished structure of music relies on the perceiver's ability to organize the individual musical events in terms of their relative stability. Hierarchies of stability exist for both single tones and for chords, although these hierarchies exhibit different properties. The hierarchy of tones is entirely context-dependent (Krumhansl, 1979; Krumhansl and Kessler, 1982; Krumhansl and Shepard, 1979), since the twelve chromatic tones are perceived as equally stable intrinsically. It is only when a certain set (a diatonic set, for instance) is selected and highlighted that a hierarchy is established. The hierarchy of chords, on the other hand, has a context-independent component and a context-dependent component, as we have shown. The chords built upon the seven steps of the diatonic scale are perceived as being intrinsically differentiated in terms of stability; additional differentiation occurs when they are perceived in the context of other chords.

It has been proposed (Krumhansl, 1979; Krumhansl et al., 1982) that stable tones and chords function as cognitive reference points, which have been described in other domains by Rosch (1975). Cognitive reference points are characterized primarily by their asymmetric relations with less stable elements. Asymmetries in psychological representations and their dependency on structural properties of the domain have been described by Rosch (1975), Tversky (1977), and Krumhansl (1978). We propose that stable tones and chords also function as cognitive reference points, and that the perception of melodic or harmonic events in relation to reference events is a crucial component of the perception of coherence in music.

Melodic structure has often been described by principles of voice-leading. One such principle is that the leading tone (one chromatic step below the tonic) is usually followed (resolved upward to) the tonic. On the other hand, a flatted second (one chromatic step above the tonic) is usually followed by (resolved downward to) the tonic. Thus, tones in the immediate pitch neighborhood of the tonic are perceived with reference to it. In general, less stable single tones are perceived with reference to more stable tones. Similarly for harmonic elements, less stable chords are perceived with reference to more stable chords. The strong asymmetries obtained in the present study bear this out. However, although the principles of voice-leading are based largely on the hierarchy of stability of tones, principles of harmonic progression do not depend as directly on the hierarchy of stability of chords. Even though the VI chord is perceived as less stable than the tonic (I), the progression from VI to I has not traditionally been considered desirable (see Piston, 1962). This is due to such factors as the overlap
of the component tones of two chords, and voice-leading considerations of the component tones. The surface structure of chord progressions may therefore not follow directly from considerations of stability. Stability is more directly related to the abstract harmonic structure rather than to the actual sequence of events that constitutes the surface string, as we will see in the section below on generative music theory.

The root functions represented in the hierarchies of tones and of chords are different (Krumhansl, 1979; Krumhansl et al., 1982; Krumhansl and Shepard, 1979). The tonic tone and the tonic chord (I) are the most stable elements in the respective hierarchies, followed by the dominant tone (perfect fifth above the tonic) and the dominant (V) chord, respectively. Beyond this, however, the hierarchies differ. The hierarchy of tones contains the third above the tonic as the remaining element of the three-element core. The hierarchy of chords contains instead the subdominant (IV) chord. This is because tones and chords have distinctive properties in their own right. At the level of single tones, pitch proximity and scale structure appear as factors. At the harmonic level, the distinctive perceptual effects of chords of different types (major, minor, diminished, and so on) as determined by diatonic structure are evident. Krumhansl and Kessler (1982) have further investigated the perceived functional relationships of single tones and chords to abstract tonal centers or keys.

**Generative music theory**

The perception of a sequence of musical events seems to involve the extraction of the more stable events and the subordination of the less stable events to them. The composer exploits these hierarchies of stability (i.e., exploits the shape of the psychological tone-space or chord-space) within the framework of some particular musical style. In the predominant Western idiom and in many forms of Eastern music as well, beginnings and endings of segments are usually marked by stable musical events. The listener familiar with this idiom has expectations as to when unstable events should resolve to more stable events, and the composer works with these expectations.

However, harmonic stability and instability must often be gleaned from a piece of music in which no chords may actually be heard as simultaneously sounded tones. The surface structure of the piece of music may, for instance, consist of a melody without any chords explicitly present. The listener extracts an underlying harmonic structure from this surface level, forming a representation at the level of chord sequences. Music theorists routinely analyze melodic sequences in terms of their chord functions.
Thus one of the ways in which two melodies can be perceived as variations of the same theme is if they have the same underlying harmonic representation. Schenker (1935/1979, 1906/1954) developed a theory that further analyzes this level of representation in terms of its most stable chords, yielding an even more abstract (reduced) level. At the most abstract (reduced) level, only the tonic (I) chord—the most stable chord in the context key—is represented. An infinite number of surface strings of musical events would map onto the same reduced levels. Schenker's theory was thus the basis for a generative theory of tonal music.

This idea has recently been developed in the form of a grammar that assigns a reduction to a piece of tonal Western music (Jackendoff and Lerdahl, Reference Notes 2, 3; Lerdahl and Jackendoff, 1977, in press). The grammar consists of a set of rules that specify the possible and preferred reductions of a surface string of musical events to increasingly abstract harmonic levels. The rules of the grammar select events from each level that are to be carried over to the next level of reduction. This selection is made on the basis of the relative stability of the musical events. Only the most stable of a group of musical events will be carried over to the next level of reduction. Lerdahl and Jackendoff's grammar determines the relative stability of events by considering metrical, grouping, melodic and harmonic factors, all of which interact. Metrical structure is a function of the regular pattern of stressed and unstressed pulses, as specified by the time-signature; events in highly stressed positions (i.e., on strong beats) are more stable than others. Grouping structure involves figural groups such as motifs or phrases (Bamberger and Schön, Reference Note 4; Cooper and Meyer, 1960). Beginnings and endings of groups may be marked by cadences, rhythmic or melodic repetitions, or the presence of rhythmic or melodic clusters that cohere in part due to Gestalt principles (Wertheimer, 1938). Each level of reduction is segmented into figural groups so that the most stable element in each group is carried over to the next level. Among the many melodic and harmonic factors influencing overall stability are whether the melodic note of a chord is in the same pitch-class as its root, whether a triad is in root position or is an inversion, and most important for our present purpose, how stable a tone or chord is relative to the other tones and chords. Lerdahl and Jackendoff's grammar thus presupposes that the listener has available a representation of the relative stability of tones and of chords in the context of the prevailing key. Schenker's system presupposes this as well, since one can know which events of a given level are to be represented at a more abstract level only if one has knowledge of the hierarchies of stability of tones and chords in the prevailing context. The hierarchy of stability of tones has been studied in previous work.
(Krumhansl, 1979; Krumhansl and Kessler, 1982; Krumhansl and Shepard, 1979). In the present paper and in Krumhansl et al. (1982), we have provided a psychological account of this hierarchy of stability for chords. At the surface level one usually does not find a sequence of chords per se. There may simply be ‘broken’ chords or melodic lines which contain numerous non-harmonic tones (appoggiaturas, passing tones, auxiliary tones, suspensions, etc.) At the very first level, the subordination of certain less stable tones to more stable ones may be a result of a process that employs the listener’s representation of the hierarchies of stability of single tones. Through stripping away melodic embellishments on this basis, the remaining tones may then be fused into abstract harmonic events by a ‘fusion’ principle postulated by Lerdahl and Jackendoff. Once this level of abstraction has been achieved, the relative stabilities of chord functions investigated here determine how the more reduced levels are achieved.

Outstanding Issues

Generative theories of language have made claims about the innateness and domain specificity of the cognitive structures necessary for the processing of language (Chomsky, 1965, 1975; Fodor, 1975). Chomsky (1975, 1980) has argued for the existence of specialized faculties for domains other than language. He suggests that the investigation of specific psychological principles that operate in specialized faculties is a more fruitful research strategy than the attempt to find general cognitive principles that operate homogeneously. Let us briefly consider which aspects of music processing may or may not be domain specific, and discuss the applicability of domain specificity and innateness arguments that have previously been given for generative theories of language.

The representation of metrical and grouping structure may be a special case of the representation of temporal organization in general, as suggested by Jackendoff and Lerdahl (Reference Note 3). They use hierarchical trees to represent one kind of reduction, time-span reduction, in which metrical and grouping factors are primarily responsible for determining the first few levels of reduction. Each branch in the tree either terminates at a longer branch or is the termination point of a shorter branch. In other words, each musical event at a given level of reduction either dominates or is dominated by its neighbor in the same figural group. Jackendoff and Lerdahl (Reference Note 3) point out that these tree structures are almost identical in form to the metrical trees in recent theories of prosodic structure proposed by Liberman and Prince (1977) and Selkirk (1980), which account
for stress and linguistic rhythm in language. They suggest that such hierar-
chical trees may describe the cognitive organization of temporal sequences
in general; certain events are represented as most salient or stable, and others
are seen as elaborations of these. Furthermore, the grouping structure in
time-span reduction may involve general principles of temporal grouping,
as outlined by Bamberger and Schön (Reference Note 4), Cooper and Meyer
(1960) and the Gestalt psychologists (e.g., Wertheimer, 1938). Metrical
and grouping principles of musical organization, at least as they have been
characterized by available accounts, seem to have commonalities with
other domains.

Melodic form can perhaps be described by general constraints on the
representation of sequential patterns of symbols (Restle, 1970, 1972, 1973;
Simon, 1972; Simon and Kotovsky, 1963) which have been applied to
musical sequences (Deutsch, 1980; Deutsch and Feroe, 1981; Simon and
Sumner, 1968). Additionally, Gestalt principles of proximity and good
continuation account for many aspects of melodic structure. For example,
small intervals predominate in melodic sequences (Jeffries, 1974; Ortmann,
1926), and melodies with larger intervals are more difficult to process and
remember (Deutsch, 1972b, 1978; Divenyi and Hirsh, 1978). Other studies
(Divenyi and Hirsh, 1974, 1975, 1978; Idson and Massaro, 1976, 1978;
Nickerson and Freeman, 1974; Idson, Reference Note 5) show better
recognition performance for simple contour patterns, that is, those with
fewer changes in pitch direction. Moreover, Dowling (1978) found that
listeners can detect contour changes even when alterations in the exact
intervals go undetected. The importance of contour has been stressed also
by Dowling (Dowling and Fujitani, 1971; Dowling and Hollombe, 1977),
Idson and Massaro (1978), and Massaro et al. (1978). Therefore, many of
the melodic principles of musical organization seem not to be specific to
music.

We have already shown how stable tones and chords may function as
cognitive reference points, which exist in other domains; thus the general
properties of cognitive reference points are not domain specific. However,
the temporal juxtaposition of tones and chords differentiated on the basis
of stability seems to be confined to music. Metrical stress, grouping, prox-
imity, good continuation and contour may exist as dimensions in various
domains, but the relationships between tones and between chords by virtue
of which they are perceived as differentially stable, independent of the
above factors, seem to be found only in music. However, to identify prin-
ciples that seem to be unique to the representation of music is to say no
more than that there are some properties of music that are not found in
other domains and that our internal representation reflects these properties.
In and of itself, this is a weak, if not trivial, sense in which a cognitive structure may be said to be domain specific.

Strong claims about domain specificity are often coupled with claims about innateness. In the linguistic domain, it is often argued that the stimulus history of the subject is too impoverished to account for the development of such complex cognitive structures through a learning process (Chomsky, 1965, 1975). Native speakers are capable of producing or understanding an infinite number of novel sentences whose surface structures are sufficiently complex that only by applying various transformations can we account for their comprehensibility. These transformations are often such that they could not plausibly be learned solely from the set of sentences to which we have been exposed.

Various discussions relating aspects of language to music have appeared since the emergence of generative linguistics (Bernstein, 1976; Jackendoff, 1977; Jackendoff and Lerdahl, Reference Notes 2, 3; Lerdahl and Jackendoff, 1977, in press; Roads, 1979; Sundberg and Lindblom, 1976; Winograd, 1968). Lerdahl and Jackendoff, for instance, suggest that claims about the innateness of musical competence follow from their generative approach. They suggest that the general form of the rules of their grammar are innate, but that the notions of relative consonance and dissonance (which are viewed here as hierarchies of stability) in their formulation may be culturally specific. Although this claim seems intuitively reasonable, capturing something general and universal about music, it is a much weaker plausibility argument than was made by Chomsky for language. Sentences in a language are, for the most part, either grammatical or not. Whereas for music, as Lerdahl and Jackendoff point out, the well-formedness constraints on their reduction trees leave open a range of possible analyses that may be assigned to a segment of tonal music, i.e., there is usually considerable structural ambiguity. This ambiguity is not only permissible in music, unlike language, but provides the basis for artistic creativity. Within the range of possible structural analyses, there are some that are more typically assigned by the listener, or more ‘preferred’. The grammar does not specify correct hearings, only preferred hearings. It is thus more difficult to evaluate the listener’s intuitions regarding the ‘grammaticality’ of musical segments.

Cross-cultural, developmental and neurological research may possibly shed some light on the innateness issue, though as yet little is known that is conclusive for music. Most cultures do seem to build their music by systematically differentiating musical elements in terms of stability. In fact, some of the specifics of the hierarchy of stability of tones of North Indian music are the same as for Western music: the tanpura drone that accompanies singers and instrumentalists consists of the tonal center and the
perfect fifth above. But there is little consensus concerning the universality of such patterns, and the cross-cultural issue often turns on what is to count as music. Developmental studies so far only indicate that very young children in experimental tasks show little evidence of differentiating tonal and atonal sequences (see, for example, Zenatti (1969), Dowling (1982) and Pick (1979) for recent reviews of the developmental literature). Whether tonal systems have not yet been learned by very young children or are simply latent remains unclear. Some neurological studies using dichotic listening tasks (Bever and Chiarello, 1974) and brain damaged patients (Shapiro et al., 1981) seem to yield fairly systematic patterns regarding the localization of musical function under various conditions. However, Bradshaw and Nettleton (1981) argue that in addition to level of expertise and talent, specific task demands seem to influence the localization of function.

Finally, it may be suggestive that although Western music has changed in many ways (such as timbre), the underlying harmonic structure has remained fairly resistant to change in music that has had wide popular appeal. Dowling (1978) claims that ‘tonal scales constitute one of the most durable families of perceptual-motor schemata that have been observed in psychology’ (p. 345). Whether this reflects an innate predisposition or simply a highly overlearned cognitive process resistant to change is unclear. In either case, we have shown in this article that the internal representation of harmony in tonal music takes a highly structured form.

References


The representation of harmonic structure in music


**Reference Notes**


**Résumé**


Les accords stables harmoniquement semblent fonctionner comme points de référence du système. On discute l’importance des représentations des hiérarchies de stabilité harmonique en relation avec des propositions de type génératif de la compétence musicale.