Tonality and expectation

1 author:

Jamshed Bharucha
Harvard University

MUSICAL PERCEPTIONS

edited by

RITA AIELLO

with

JOHN A. SLOBODA

New York *Oxford*

OXFORD UNIVERSITY PRESS

1994

9

Tonality and Expectation

JAMSHED J. BHARUCHA

Introduction

Jamshed Bharucha's research has been highly influential in proposing that tonality gives rise to both expectations and aesthetic experiences (Bharucha, 1984). In this chapter Bharucha discusses how musical expectations are generated experimentally. He describes the relationship between tonal expectations, consonance, stability, and memory, and questions whether musical expectations are learned, innate, or based on the structure of the sounds themselves. His interest is in understanding the cognitive processes that underline tonal expectations and their acquisition. Bharucha points out that when listening to music listeners rejoice at what they already know; therefore, he asks, are they giving up the element of surprise or the element of familiarity? For Bharucha, tonal expectations are a compelling aspect of the aesthetic experience of listening to music. He divides expectations into two kinds: schematic expectations, which are culturally generic, automatic, and related to a general schema, and veridical expectations, which are related to the actual piece of music, to a specific memory.

Bharucha explains how he has used neural networks to account for certain findings about music perception. Neural networks are computer models that simulate brain-like systems based on fundamental principles of neural organization: they have been used to model human behavior. The application of neural networks to the perception of music is a promising and fast-growing area of research (see Bharucha, 1987 and 1991; Gjerdingen, 1989; Todd & Loy, 1991).

REFERENCES

Bharucha, J. J. (1984). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology, 16,* 485–518.

Bharucha, J. J. (1987a). MUSACT: A connectionist model of musical harmony. In *Ninth Annual Conference of the Cognitive Science Society* (pp. 508–17). Hillsdale, NJ: Erlbaum.

Bharucha, J. J. (1987b). Music cognition and perceptual facilitation. A connectionist framwork. *Music Perception, 5,* 1–30.

Bharucha, J. J. (1991). Tonality and learnability. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Gjerdingen, R. O. (1989). Using connectionist models to explore complex musical patterns. *Computer Music Journal, 13*(3), 67–75.

Todd, P. M., & Loy, G. D. (Eds.) (1991). *Music and connectionism.* Cambridge, MA: MIT Press.

Why does a segment of music ending in a dominant chord generate strong expectations for the tonic to follow? What is the relationship between tonal expectations, consonance, stability, and memory? Are expectations of this kind learned, innate, or based on the structure of sound? If they are learned, how?

Some of these and related questions about tonality and expectation have been pondered through the ages by music theorists,[1] and today we benefit from a wealth of work on which to build. In this chapter I shall sketch an approach to these questions based on psychological experiments and computational modeling designed to explore the cognitive processes underlying tonal expectations,[2] their relationship to consonance, stability, and memory, and their acquisition.[3]

Tonal expectations are important to study for several reasons, including their role in the aesthetic experience of music, their role in the mental organization of music, their influence on consonance and dissonance, and their role in learning. The advantage of viewing these disparate phenomena in terms of expectation is that there is the promise that the processes that underlie expectations underlie aspects of these other phenomena as well.

The interested reader is directed to a substantial body of important work that will not be reviewed in this chapter, particularly work on the specific musical determinants of expectation (Meyer, 1956; Carlsen, Divenyi, & Taylor, 1970; Narmour, 1977, 1990; Carlsen, 1981; Schmuckler, 1989). The focus in this chapter will be on trying to understand the perceptual processes that underlie

REFERENCES

Bharucha, J. J. (1984). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology, 16,* 485–518.

Bharucha, J. J. (1987a). MUSACT: A connectionist model of musical harmony. In *Ninth Annual Conference of the Cognitive Science Society* (pp. 508–17). Hillsdale, NJ: Erlbaum.

Bharucha, J. J. (1987b). Music cognition and perceptual facilitation. A connectionist framwork. *Music Perception, 5,* 1–30.

Bharucha, J. J. (1991). Tonality and learnability. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Gjerdingen, R. O. (1989). Using connectionist models to explore complex musical patterns. *Computer Music Journal, 13*(3), 67–75.

Todd, P. M., & Loy, G. D. (Eds.) (1991). *Music and connectionism.* Cambridge, MA: MIT Press.

Why does a segment of music ending in a dominant chord generate strong expectations for the tonic to follow? What is the relationship between tonal expectations, consonance, stability, and memory? Are expectations of this kind learned, innate, or based on the structure of sound? If they are learned, how?

Some of these and related questions about tonality and expectation have been pondered through the ages by music theorists,[1] and today we benefit from a wealth of work on which to build. In this chapter I shall sketch an approach to these questions based on psychological experiments and computational modeling designed to explore the cognitive processes underlying tonal expectations,[2] their relationship to consonance, stability, and memory, and their acquisition.[3]

Tonal expectations are important to study for several reasons, including their role in the aesthetic experience of music, their role in the mental organization of music, their influence on consonance and dissonance, and their role in learning. The advantage of viewing these disparate phenomena in terms of expectation is that there is the promise that the processes that underlie expectations underlie aspects of these other phenomena as well.

The interested reader is directed to a substantial body of important work that will not be reviewed in this chapter, particularly work on the specific musical determinants of expectation (Meyer, 1956; Carlsen, Divenyi, & Taylor, 1970; Narmour, 1977, 1990; Carlsen, 1981; Schmuckler, 1989). The focus in this chapter will be on trying to understand the perceptual processes that underlie expectation. Toward this end, only a small set of specific musical expectations will be addressed—tonal harmonic expectations—by way of illustrating the potential of certain experimental and computational approaches. The reader is also directed to the work of Jones (1981a, 1981b) for a more overarching theoretical approach to understanding musical expectation.

Tonal Expectation and Aesthetics

Tonal expectations are a compelling aspect of our aesthetic experience of music. The deception of a deceptive cadence, the irresolution of a half cadence, and the finality of a final cadence are only the most obvious examples of tonal expectations at work in Western tonal music. Delayed resolution of tonal expectations is often a central distinguishing characteristic ascribed to late romantic composers such as Wagner. Tonal expectations can be found in other cultures as well. Some of the most dramatic moments of a performance of Indian classical music involve an eventual resolution to the tonal center after a prolonged yearning for it.

The most influential modern exponent of the role of expectation in musical aesthetics is Leonard Meyer (1956). According to a widely accepted view based largely on the work of Meyer, a piece of music in a familiar genre generates expectations, the subtle violation of which is emotionally important. There is considerable support in psychology for the relationship between expectancy violation and emotion (e.g., Mandler, 1984).[4]

Little, if any, work has been done by cognitive psychologists on the affective aspect of experiences associated with expectation. In contrast, a fair amount of research has been done on the processing of sound that leads up to (and is a prerequisite to) this affective experience. This chapter thus focuses exclusively on trying to understand these prerequisite processes, without taking a position on the nature or interpretation of the affective experience induced thereby.

The Puzzle of Expectations

There is a puzzle concerning the role of expectations in music that Dowling and Harwood (1986) call "Wittgenstein's paradox": *The puzzle is that when we are familiar* with a piece of music, there can be no more surprises. Hence if expectancy violation is aesthetically important, a piece would lose this quality as it becomes familiar.

Attempts to resolve this apparent paradox often postulate a mental structure, called a schema (Dowling & Harwood, 1986), that has assimilated the music of a genre over a lifetime of experience. The expectations produced by this system are generic in that they represent the most likely transitions in one's musical culture.[5] They are also automatic, in the sense that they cannot easily be suppressed. Even when a given piece has been heard often enough to be familiar, it cannot completely override the generic, automatic expectations. Surprises in a new piece thus continue to have a surprising quality because they are heard as surprises relative to these irrepressible expectations.

This line of reasoning seems necessary but cannot be sufficient. If the surprises in a new piece continue to be surprises even after repeated hearing, the piece would never sound familiar. Furthermore, a performer playing from a schematic representation would be unable to predict the next event, but would instead drift into the most generic progression; the performer would, for example, always follow a dominant chord with a tonic. Using this explanation, you either give up surprise or you give up familiarity.

Yet it clearly must be the case that an atypical transition in a familiar piece is both familiar and surprising. It is familiar in a sense and surprising in another. It is thus necessary to postulate two kinds of expectations, which may be called *schematic* and *veridical* (Bharucha, 1987a). Schematic expectations are the automatic, culturally generic expectations mentioned above. They must be generated by a system that has learned to expect the events that are the most likely. Veridical expectations are for the actual next event in a familiar piece, even though this next event may be schematically unexpected. They must be generated by a system that has learned that particular piece. Since schematic expectancies are acquired from hearing many individual pieces, the two kinds of expectancies will converge more often than not for pieces typical of one's musical culture. Sometimes they will diverge, however, creating the sense of violation of which Meyer wrote.

Expectation and Mental Representation

The existence of two kinds of expectations suggests (but doesn't require) the existence of two systems that generate them. These may be called a *schema* and a *memory*, which generate schematic and veridical expectancies, respectively. A schema is a term used by cognitive psychologists to refer to a mental representation of generic relationships. The memory, in contrast, would represent specific sequences of music in a way that enables them to be retrieved (recognized or recalled) individually. A schema enables one to recognize the style of a new piece; a memory enables one to recognize a particular piece.[6]

A schema generates expectations for events (such as chords) that *typically* follow. A schema thus embodies the information that would generate *typical* hierarchies of stability. These hierarchies, called *tonal hierarchies*, reflect the relative stability of tones or chords in given key-inducing contexts.

A memory, in contrast, generates expectations for events that *actually* follow in a particular familiar piece of music. A memory thus embodies the information that would enable a listener to recognize an error in a familiar piece, and that would enable a performer to play the correct note, however schematically unexpected.

Expected events are heard as more consonant and more stable than unexpected events, as will be seen below. Thus expectation, consonance, and stability covary with the musical context. If this is true, these terms refer to slightly different experiential aspects of the same underlying process.

The activation of a schema while listening means that, in addition to hearing the sounded pitch events (along with their temporal, timbral, and other qualities), schematic expectations are generated and are fulfilled or violated to varying degrees, with concomitant effects on stability. The schema thus imposes a cultural elaboration on the sounded events. Elaboration is widespread in perception and cognition (see Anderson, 1990). While reading a story, for example, inferences are drawn and familiar themes are triggered. These happen automatically, unavoidably, and constitute a mental elaboration of the text.

The representations postulated by Lerdahl and Jackendoff (1983) and by Deutsch and Feroe (1981) are essentially representations of mentally elaborated pieces of music. According to Lerdahl and Jackendoff, a piece of tonal music is represented as an inverted tree, of which the pitch-time events of the piece are at the terminal branches (see Figure 9.1). An inverted tree diagram shows levels of abstraction from the piece of music (eighth-note level, quarter-note level, etc.).

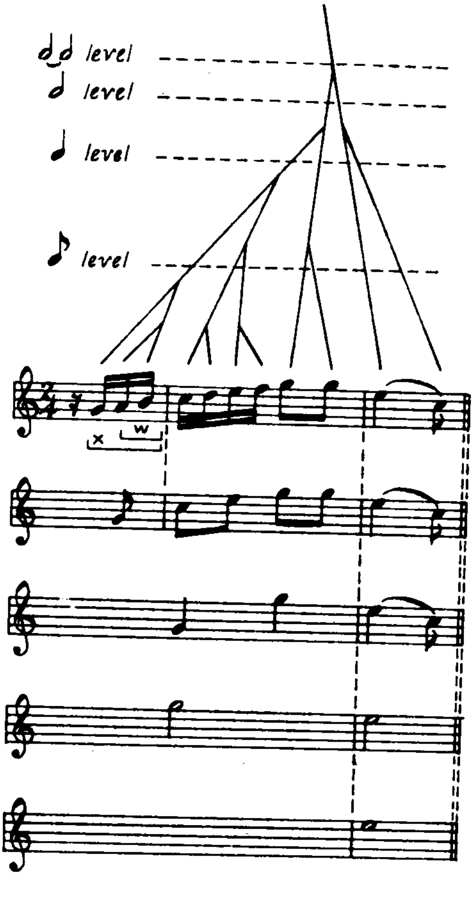

FIGURE 9.1. An inverted tree diagram shows levels of abstraction from the piece of music (eighth-note level, quarter-note level, etc.). A branch that terminates at a longer branch represents the subordination of an event to a more stable event. The most stable event (E in the last measure) is represented by the uninterrupted trunk of the tree. The increasing levels of abstraction are written out in musical notation below the score of the piece. (Reprinted from Lerdahl & Jackendoff, 1983. Copyright © 1983 by M.I.T. Press.)

A branch that terminates at a longer branch represents the subordination of an event to a more stable event. The most stable event (E in the last measure) is represented by the uninterrupted trunk of the tree.

A piece of music is thus represented simultaneously at various levels of abstraction. We may call this an *event hierarchy* (see Bharucha, 1984). An event hierarchy specifies how the events in a particular piece are organized hierarchically. An event hierarchy is a hierarchy of specific pitch-time events (e.g., a particular C in this piece dominates the D passing tone following it; see Figure 9.1). In contrast, a tonal hierarchy is a hierarchy of pitch classes (e.g., when the key of C major has been established, C's are in general more stable than D's). An event hierarchy is a mentally elaborated representation of the piece of music.

Lerdahl and Jackendoff specify a set of rules for determining which branches fuse into larger branches, and which of two fusing branches dominates the other. One of their rules determines which of two branches dominates based on which is more stable. How does the listener know which is more stable? This is precisely what the tonal hierarchy specifies. An event hierarchy thus presupposes a tonal hierarchy. The tonal hierarchy, generated continuously by the schema, is the cultural elaboration that is imposed on the music, thereby enabling it to be heard as an event hierarchy.

In the areas of visual perception and language, a debate has raged among cognitive psychologists as to whether events are stored in memory in the pristine form in which they occur or whether the elaboration is stored together with them. For example, are the inferences and themes that are automatically triggered while reading a story stored along with the words themselves? In the case of stories, there seems little doubt that aspects of the elaboration are stored; in fact, it's more likely that the reader will be able to recall the theme than the text of the story itself.

The evidence from cognitive domains other than music thus strongly suggests that a piece of music is remembered not just as a recording of the sound but in all its mentally elaborated glory.[7] Storing the entire event hierarchy would make it possible to retrieve selected levels. In this way, one could recall the piece at abstract levels of the tree without having to play it back in one's head, note by note. This possibility seems congruent with casual experience.

The following is a summary of experimental and computational work that supports aspects of the theoretical framework sketched here—in particular, schematic expectations, their relationship to consonance and stability, and an exploration of how they might be learned.

The Measurement of Schematic Expectations

Tonal expectations have been studied quite exhaustively, both theoretically (Narmour, 1977, 1990) and empirically (Carlsen, Divenyi, & Taylor, 1970; Carlsen, 1981; Schmuckler, 1989). The experiments discussed below employ a method called priming. They were designed to measure schematic expectations in such a way as to provide some insight into the underlying processes. They were also designed to be able to detect schematic expectancies in the minds of listeners with no formal musical training whatsoever.

In the priming task, listeners hear a musical sample (called the prime) followed by a chord (called the target chord). If the target is an expected chord, it should be processed more quickly than if it is an unexpected chord. Speed of processing is a widely accepted measure of expectancy in cognitive psychology. This is because expectation is thought to play a preparatory role. If something is expected, the perceptual system prepares for it; if that expected event occurs, the advance preparation results in faster processing. This is an adaptive strategy because it enables us to anticipate events in our environment.

In order to measure the speed with which the target is processed, it is essential to have the listener make some binary judgment about it and then press a button indicating the judgment. A computer records which button was pressed and the elapsed time from the onset of the target to the button press (called the reaction time). Since we are interested in the time it takes to process the target in a musically meaningful way (as opposed to the time to simply detect its presence), the binary judgment must be one that focuses the listener's attention on a musically relevant aspect of the target. The judgment we chose was an intonation judgment. On half the trials, the target was mistuned by flattening the fifth degree of the chord by approximately 1/4 of a semitone. The listener's task is therefore to press one designated button if the target is in-tune and another if it is out-of-tune.

In order to remove all subjectivity from the intonation judgment, the main experiment is preceded by a session in which target chords are presented in the absence of primes, and listeners are given feedback as to the correctness or incorrectness of their response. Thus the criterion for what counts as in-tune and what counts as out-of-tune is established objectively for all subjects. This last point is critical in order to study listeners who have had no formal musical training. When this group of listeners is studied in experiments, their responses often show very little clear structure with open-ended judgments (such as "rate the intonation on a scale from 1 to 7"), and the subjects express considerable uncertainty about what they are doing, because they are inexperienced at making such judgments. When the judgment is binary and there is a correct and incorrect answer, however, these subjects understand the task immediately and reveal clear patterns in their responses.

A binary true/false judgment is also useful for studying professional composers or theorists. This group tends to be understandably skeptical about the validity of making such simplistic judgments as are used in psychological experiments, and also finds that they can't help making the judgments based on their analytical knowledge. But when there are clearly defined correct and incorrect answers (and when they are persuaded by the experimenter that "correct" and "incorrect" do not imply value judgments), they abandon their theoretical qualms and seek to maximize the percent of trials they get correct (which is displayed at the end of the experiment).

Another reason for employing an intonation judgment is that it enables us to study how the perceived consonance of a chord is influenced on the basis of whether or not it was expected. To summarize, the priming task using an intonation judgment has two goals: (1) to measure expectation using reaction time, and (2) to study the effect of expectation on consonance. The task is designed so as to pursue these goals for musically untrained as well as trained listeners.

From trial to trial, the experimenter varies the degree to which the target chord is expected[8] given the prime. In the simplest such experiment (Bharucha & Stoeckig, 1986, experiments 2 and 3), the prime was a single chord, and the target was either most highly expected or most highly unexpected given the prime. Thus if the prime was C major, the target was either G major or F-sharp major, as shown in Figure 9.2 (all keys were used in the experiment).[9]

This yielded four types of trials: in-tune expected, out-of-tune expected, in-tune unexpected, and out-of-tune unexpected. The listeners only needed to concern themselves explicitly with whether or not the target was in-tune or out-of-tune.

The results were as follows. First, listeners were faster and more accurate, on average, to make intonation judgments for the expected chords than for the unexpected chords (see table). This demonstrates that the process underlying expectation is one of preparation or facilitation. It is as if a representation of the expected chord had been mentally activated even before the chord occurred, based on its being highly probable in the familiar tonal style.

The second central result was that listeners were faster and more likely to respond "in-tune" when the chord was expected than when it was unexpected, or, equivalently, they were faster and more likely to respond "out-of-tune" when the chord was unexpected than when it was expected. This indicates that expected chords sounded more consonant, or, equivalently, unexpected chords sounded more dissonant. This result shows starkly that there are at least two determinants of consonance. One is the relationship between simultaneous sounds. This in turn is based on the way sound interacts with the inner ear (see below) as well as learning. The second determinant is the relationship between

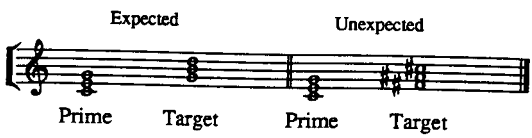

FIGURE 9.2. The design of the priming experiment (Bharucha & Stoeckig, 1986). The figure shows root position triads in close position only to illustrate the difference between expected and unexpected. The actual chords heard by the subjects were considerably more complex (see note 9).

Results of priming experiment
(in milliseconds)

	Expected	*Unexpected*
In-tune	788	982
Out-of-tune	864	803
Average	826	893

two sequential sounds. This, it is claimed here, is primarily learned and is mediated by expectations.

Both effects described above were observed for rank novices and professionals alike. The novices were novices only insofar as they had no formal training and no experience playing an instrument. But by virtue of having grown up in a particular culture (in this case, the United States), they have presumably learned what to expect from extended passive exposure.

It is important to note that expected events aren't necessarily preferred over unexpected events. This caution about expectation holds true of consonance and stability as well. If there is any relationship between expectation and its associated concepts, on the one hand, and preference, on the other, it takes the form of an inverted-U function (see Dowling & Harwood, 1986, for a discussion): A moderate amount of violation of expectations is generally preferred over always fulfilling expectations or always violating them.

Even this trend should be viewed with some caution. People who are immersed in experimenting with new musical forms may develop preferences based on phenomena other than expectation, consonance, and stability. Although a composer may have exposure to several genres, including some that are very different from tonal-harmonic music, and would thus develop a rather diverse set of schemata, it is almost impossible to escape exposure to tonal-harmonic music if one grows up in a Western society. Tonal-harmonic music pervades television, radio, supermarkets, and the dominant popular forms of music. Some of the subjects in the above priming experiment were composers who viewed the experiment with grave suspicion, seriously doubted they would show any evidence of schematic tonal expectations, and asserted their lack of preference for expected tonal events. Nevertheless, they showed the same pattern of results as novices. A person is in a sense a prisoner of the schematic expectations driven by the inadvertent learning of musical patterns that pervade the environment.

This means that it is difficult to suppress schematic expectations. Can this be demonstrated empirically? In one experiment, 80% of the trials contained highly unexpected chords and only 20% contained highly expected chords. Even over a session of 100 trials, the chords that were processed more quickly were those that were traditionally expected even though they were in the clear minority during the session (Bharucha & Stoeckig, 1989). Distorting the musical environment in this way was not sufficient to wipe out a lifetime of exposure.

In another experiment, each prime-target pair was played twice in succession. For example, in an "unexpected" trial, the subject would hear C-F-sharp-pause-C-F-sharp. The last chord was mistuned on half the trials, and subjects had to decide whether it was in-tune or out-of-tune. The last chord, F-sharp, is schematically unexpected (since it rarely follows C) but veridically expected (since it actually follows C and the subject knows the same pair will be repeated). Nevertheless, schematically expected chords are still processed more quickly than schematically unexpected chords (Bharucha & Stoeckig, 1989). These experiments thus provide some preliminary evidence that schematic expectations are very difficult to suppress.

Expectation and Tonal Hierarchies

There has been a more extensive study to date of tonal hierarchies than of expectations. Since one of the claims in this chapter is that the processes that underlie schematic expectations may also be responsible for tonal hierarchies, a summary of research on tonal hierarchies is relevant.

Tonal hierarchies have been studied most exhaustively by Krumhansl (1979, 1990; Krumhansl & Shepard, 1979). The term "tonal hierarchy" signifies the relative differences in stability of tonal elements. For example, when the key of C major has been strongly established during a piece of music, the diatonic tones (C, D, E, F, G, A, B) are more stable than the nondiatonic tones. If a different key is established, a different set of tones will be the stable diatonic set. Among the diatonic set in any given key, the tonic is the most stable, and the major third and perfect fifth are generally more stable than the remaining four.

Although this tonal hierarchy for tones in the context of an established key is well known from elementary music theory of the common practice period, Krumhansl sought to quantify this hierarchy. Among the many advantages of quantifying a tonal hierarchy are the following. First, although the tonal hierarchy is quite clear when a key is strongly and unambiguously established, it is less clear when a key is weakly established or when there is more than one established key (as in polytonality). A quantification of tonal hierarchies makes it possible to track the development of keys even in the latter case. For example, there has been considerable debate as to whether the opening passage of Stravinsky's *Petroushka* is simultaneously in C major and F-sharp major or, alternatively, in no key at all. By quantifying the tonal hierarchy in this passage, Krumhansl and Schmuckler (1986) were able to determine that indeed both keys, C major and F-sharp major, were established.

A second advantage (perhaps the more significant psychologically) of quantifying the tonal hierarchies is that it makes it possible to depict tonal relationships spatially. For example, the circle of fifths is a spatial depiction of the relationships between major keys (see Figure 9.3). A single spatial representation such as a circle captures, all at once, the many pairwise relationships between

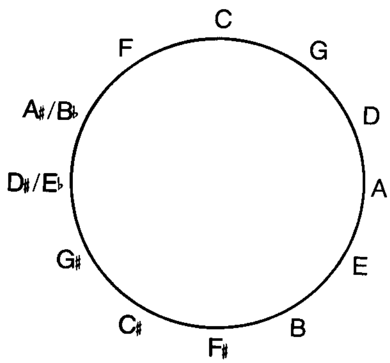

FIGURE 9.3. The circle of fifths.

the 12 possible major keys. Thus, instead of individually enumerating the relationship between all the keys (e.g., C is most closely related to G and F, less closely related to D and B-flat and so on, G is most closely related to D and C, less closely related to A and F, and so on, and so forth), a single diagram with the 12 keys labeled shows all the relationships simultaneously. The advantage is even more striking if minor keys are included. A single spatial representation that captures all these pairwise relationships at once provides a vivid insight into the psychological representation of tonality. That is not to say that the same spatial representation necessarily exists physically in the brain.[10] Rather, it means that the physical representation in the brain encodes the information that is contained in the spatial representation.[11]

At this point one might be inclined to ask why quantification of tonal hierarchies is an advance, since music theorists have a long tradition of tracking keys and employing spatial representations (the circle of fifths, in particular). The answer is that methods of quantification permit us to investigate tonal hierarchies in the minds of listeners other than theorists, who, after all, have extensive formal training and who, by the very nature of the enterprise, adopt a highly analytical stance toward listening.

The method Krumhansl has used most extensively to investigate tonal hierarchies is called the *probe tone technique*, originally developed with Roger Shepard (Krumhansl & Shepard, 1979). In this technique subjects hear a musical context followed by a probe tone. The musical context could be either a musical scale, a chord, a chord progression, or an actual sample of music. The *probe tone* is picked at random from the 12 chromatic pitch classes. The listener's task is to rate how well the probe tone fits within the musical context by rating it on a scale from 1 to 7. The context is then repeated and a different probe tone is presented. The context has been fully probed when all 12 chromatic pitch classes have been presented. This process may be repeated for the same listener and also repeated for many listeners, yielding several replications of each of the 12 probe tones. The ratings for each probe tone are averaged across the replications, yielding 12 numbers, one for each probe tone. This set of 12 numbers is called a *profile* (see Figure 9.4).

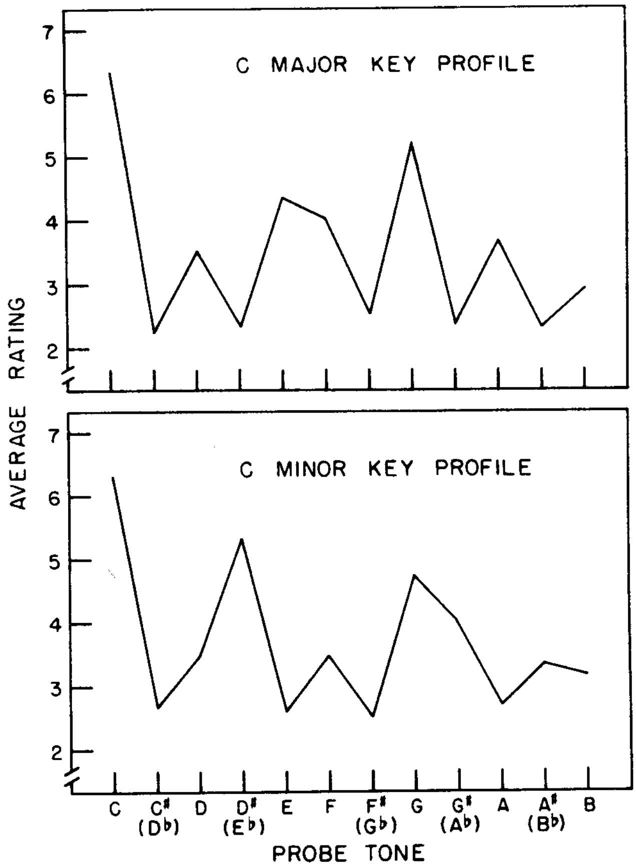

FIGURE 9.4. Profiles for major and minor keys. The rating for any given probe tone depicts the perceived stability of that tone in that key. The profiles are shown with respect to C major and minor for convenience only. (Reprinted from Krumhansl & Kessler, 1982. Copyright © 1982 by the American Psychological Association. Reprinted by permission.)

When the musical sample establishes a key strongly and unambiguously (e.g., a IV-V-I cadence), the profile shows the tonic being the most highly rated, followed by the other two tones of the tonic triad, followed by the remaining diatonic tones (Krumhansl, 1979, 1990). The profiles from samples with a strong and unambiguous key are averaged and the resulting profile is used as a standard profile for that key. The standard profiles for the 24 major and minor keys constitute the standards against which the profiles of samples with weaker or more ambiguous keys are assessed.

Krumhansl and Kessler (1982) showed how one can track the development of keys in a musical sample of interest. To do this, the profile obtained for that sample is correlated with each of the 24 standard profiles. The strength of the sample's correlation[12] with a key standard indicates the strength of that key in the sample as of the moment at which the probe was presented. By probing a piece of music at different points in time, the development and change of keys over time can be tracked.

To represent the relationships between keys spatially, the standard profiles for the 12 major and minor keys are themselves correlated with each other. The higher the correlation between two keys, the closer they are in the spatial representation. Thus C major is highly correlated with G major and F major, as well as with A minor and C minor, and therefore needs to be located close to them. In addition, it has a range of correlations with the other keys, which must therefore be located at corresponding distances. Simultaneously satisfying all these constraints is very difficult. This is accomplished by a computerized method called multidimensional scaling, developed by Shepard (1964). Krumhansl and Kessler (1982) found that the spatial representation that best satisfies these multiple constraints is a torus. A torus requires four dimensions, but for convenience can be imagined in three dimensions as a doughnut, with the proviso that the inner and outer perimeters of the doughnut be equal (which is why three dimensions don't suffice). This solution captures the circle of fifths for major keys and for minor keys, and puts the two together in such as way as to preserve the relative and parallel major-minor relationships. The relative positions of keys on the torus can be examined by cutting it and unfolding it into a flat sheet, as shown in Figure 9.5. A torus results from joining the top and bottom edges

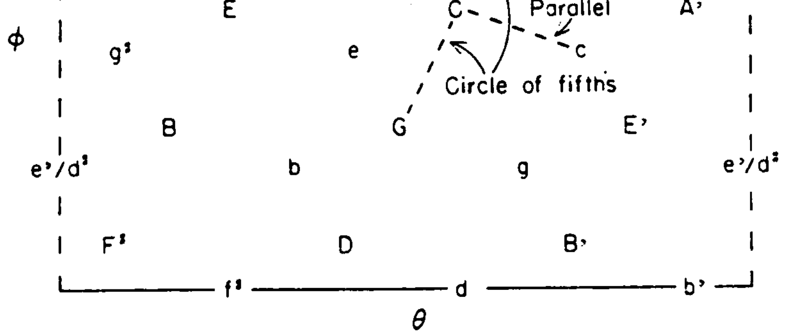

FIGURE 9.5. An equivalent two-dimensional map of the multidimensional scaling solution of the 24 major and minor keys. (Reprinted from Krumhansl & Kessler, 1982. Copyright © 1982 by the American Psychological Association. Reprinted by permission.)

together and joining the left and right edges together. If you actually try this with a piece of paper, it will either wrinkle or rip, which is why an additional dimension is needed.[13]

Tonal hierarchies usually, but not always, reflect the frequency of occurrence of tones just heard: tones heard more often are judged to be more stable. When American college students were tested with samples of North Indian music, for example, their profiles were very similar to those of Indian students (Castellano, Bharucha, & Krumhansl, 1984). The profiles of both groups of listeners matched the hierarchies described by theorists of Indian music (Jairazbhoy, 1971), but they also reflected the frequency of occurrence of tones in the samples. This should not be surprising: Tones that are heard as stable have typically occurred more often in the preceding sample. Numerous studies of note counts (summarized extensively by Krumhansl, 1990) show this to be true. In tonal music of the common practice period, tonics tend to occur most often, followed by the other tones of the tonic triad, and then the remaining diatonic tones. Tonal pieces tend to reveal their tonality in the selection of pitch classes and their relative priority. Tonal hierarchies obtained in these experiments thus show, at the very least, that listeners have a memory system that can integrate the pitch classes heard in the immediately preceding sample (called the *proximal context*). This integrated representation then leads them to hear pitch classes that immediately follow as being differentiated on the basis of stability.

The distribution of pitch classes in the proximal context is not the sole determinant of the tonal hierarchy. The profiles of the Indian students in the study by Castellano, Bharucha, and Krumhansl (1984) were found to be influenced by an additional factor: past exposure (called the *distal context*) to these musical forms. This determination was made by a regression analysis. A regression analysis enables one to determine the extent to which different factors contribute individually to an outcome. After determining that frequency of occurrence of pitch classes in the proximal context was the strongest determinant of the profiles for both American and Indian listeners, frequency of occurrence was factored out, and the residual portions of the profiles were analyzed for the influence of additional factors. The residuals of the Indian listeners correlated with profiles based on music theory (Jairazbhoy), whereas the residuals of the American listeners did not. Since the essential difference between the two groups of listeners was their background, it can be concluded that the ratings given by the Indian listeners were based in part on their familiarity with that genre and not just on the information contained in the proximal context.

It is this residual effect of the distal context—that is, prior exposure—that justifies the postulation of a schema, a mental structure that encodes the regularities of a genre as opposed to the specifics of individual pieces. There must nevertheless be enough information in the proximal context to engage the appropriate schema; after all, we have schemas for many things, possibly even many different musical genres. The information in the proximal context (such as relative occurrence of different pitch classes) serves to activate the appropriate schema, which then generates expectancies that may compensate for the absence of tones that typically occur in those contexts.

There may be other determinants of tonality besides the distribution of pitch classes in either the proximal or distal context. Among them are the rareness of the intervals (see Chapter 8 by Butler and Brown in this volume).

Can Schematic Expectation Be Explained by the Structure of Sound and the Physiology of the Ear?

Consider two tones an octave apart, each containing eight harmonics. Every harmonic of the higher tone is present in the lower tone, and every second harmonic of the lower tone is present in the higher tone. This can be represented by the ratio 1:2, which, for convenience, we shall call the *overlap ratio*. Now consider two tones a perfect fifth apart. Every second harmonic of the higher tone is present in the lower tone, and every third harmonic of the lower tone is present in the higher tone. This interval has an overlap ratio of 2:3. For two tones a perfect fourth apart, every third harmonic in the higher tone is present in the lower tone, and every fourth harmonic in the lower tone is present in the higher tone. This interval has an overlap ratio of 3:4. If you are familiar with how these intervals sound, you know that an octave is the most consonant interval, followed by the perfect fifth, followed by the perfect fourth. As the overlap ratio gets more complex, the intervals sound more dissonant.

Pythagoras, the ancient Greek mathematician and philosopher, arrived at the same ratios more than 2,000 years ago on the basis of the lengths of stretched strings. A string player produces an octave above an open string by stopping the string at half its length, and produces a perfect fifth by stopping the string at two-thirds of its length, and so on. Since Pythagoras couldn't have known much about harmonics, and even less about the workings of the ear, he concluded that there was something inherently special about simple ratios. Today the most compelling account of why consonance is related to ratio simplicity is in terms of how the harmonics affect the inner ear.

The inner ear serves as a prism for sound: it separates sound into its various frequency (pure tone) components. For example, a tone with eight harmonics will cause the basilar membrane in the inner ear to vibrate at eight different places along its length. The vibration of the basilar membrane causes hair cells that are lined up against it to respond. The response of these hair cells thus constitutes a (spatial) representation of the sound spectrum.

Two frequencies that are very close together cause the basilar membrane to vibrate at two points that are close together. Under these circumstances, the membrane's response to one interferes with its response to the other. This produces a sensation of roughness that is typically heard as dissonant. Thus the more complex the overlap ratio between two tones, the greater is the extent to which the vibrations of the basilar membrane caused by one will interfere with those caused by the other, and the more dissonant the tones will sound together.

If the consonance of intervals can be explained by spectral overlap, can other relationships in musical harmony be explained in a similar way? At first glance, this approach might seem promising. Chords that are considered most closely related tend to have tones in common—e.g., tonic-dominant, parallel major and minor, and relative major and minor relationships. In contrast, chords that are considered most unrelated have no tones in common, C and F-sharp major. Could it be that schematic expectations are simply driven by common tones?

To answer this question, the priming paradigm mentioned earlier was employed. In the original priming experiments, the expected target chord was one step away from the prime chord along the circle of fifths; for example, if the prime was C major, the target was G major. Notice that these two chords share a component tone (as well as all the harmonics of this shared tone). In contrast, the unexpected target was diametrically opposite the prime chord on the circle, and these two chords share no component tones. Thus the faster and more accurate responses to the expected chord could possibly be due simply to the fact that some of its tone components had just been heard. To eliminate this problem, it was necessary for the expected target to have no shared tones with the prime. This condition is easily met by selecting a target two steps away from the prime along the circle of fifths. For example, a C major prime leads to a greater expectation for D major than for F-sharp major, yet shares tones with neither. To strengthen the case, all harmonics that were common to the prime and target were removed, thereby ensuring that the expected target wasn't favored because of its physical similarity to the prime. In spite of this, the expected chord was processed faster and more accurately and was more likely to be judged in-tune than the unexpected chord (Bharucha & Stoeckig, 1987).

There are purely theoretical reasons to reject spectral overlap as the sole determinant of schematic expectations. Consider the following pairs of major triads: C-D versus C-A. C and D do not share any tones, whereas C and A share one tone (E). Thus C and D have less spectral overlap than do C and A. Yet C and D are considered to be more closely related to each other than are C and A, because they are closer along the circle of fifths of chords (see Figure 9.3).

Before rejecting spectral overlap as the basis for harmony, however, perhaps we should question the validity of the circle of fifths for chords. The circle of fifths is typically used to describe the relationship between keys (of the same mode). The (shortest) number of steps from one key to another along the perimeter of the circle represents exactly the number of tones present in the diatonic set of one key but not in the diatonic set of the other. For keys (of the same mode), then, spectral overlap (i.e., the intersection of the diatonic sets of different keys) exactly mirrors the circle of fifths. For chords, however, it does not. Yet the circle of fifths accurately describes the harmonic distances between chords, because chords closer on the circle are more likely to occur in the same tonal piece (they have more parent keys in common).

Since we cannot rest our case on the mere use of the circle of chords as a descriptive device, a test is needed. The appropriate test is to see whether a chord generates stronger expectations for a chord that is closer on the circle of fifths but doesn't share a tone than for a chord that is further but does share a tone. This experimental design pits convention against acoustics in determining expectations. If schematic expectations are acoustically driven, C major would generate stronger expectations for E major (with which it shares a tone) than for D major (with which it shares no tones).[14] If, on the other hand, schematic expectations are driven by a schema that has encoded the chord relationships on the basis of their conventional typicality, C major should generate stronger expectations for D major than for E major. Recent results support the latter outcome (Tekman & Bharucha, 1991).

If schematic expectations are driven by spectral overlap, then any impairment of one's ability to process auditory spectra ought to be accompanied by a corresponding diminution of schematic expectations while listening to music. A man suffered a stroke that destroyed much of the auditory cortex of his brain. Although he could detect the presence of sounds at almost normal levels, he simply could not tell the difference between the in-tune and out-of-tune chords used in the priming experiments. He was unable to detect a mistuning so severe that people with normal hearing cringe almost without exception. In the priming task, his performance was never much above 50%, which would be expected from sheer guessing. What is remarkable is that he tended to judge expected targets to be in-tune and unexpected targets to be out-of-tune (Tramo, Bharucha, & Musiek, 1990). In other words, his sense of schematic expectation was intact even though his ability to process fine-grained spectral information was severely impaired. It seems that as long as some auditory information get through, the relevant schema can be activated, thereby generating schematic expectations. This is good news for people who experience a diminished ability to process the full range of the audible spectrum, as is typical with age.[15]

Even the relative consonance of simultaneously sounded tones cannot entirely be explained by the structure of sound and the mechanical characteristics of the ear. For otherwise one should be able to synthesize artificial sounds that have very different spectra, as Pierce (1969, 1983) has done, and predict the relative consonance of intervals formed with these tones. Although there is evidence for this using tones with nonharmonic partials that are so discrepant from harmonics that the tone has no clear pitch (Geary, 1980), cases in which the tones have clear pitch reveal an overwhelming tendency to hear the intervals as we are accustomed to.

Consider a clarinet tone, which has very little energy at the even-numbered harmonics. If the fundamental frequency is 440 Hz, the next harmonic would be $440 \times 3 = 1320$ Hz. The tone that should sound most consonant with the 440-

Hz tone is one whose fundamental frequency is 1320. All the harmonics present in the higher tone would also be present in the lower tone, and every third harmonic in the lower tone would also be present in the higher tone, yielding a 1:3 ratio of spectral overlap. This yields an interval of an octave and a fifth. Two clarinet tones an octave apart have no spectral overlap at all. If spectral overlap is the sole determinant of consonance, two clarinet tones should sound more consonant when they are an octave and a fifth apart than when they are an octave apart. Indeed, many other intervals should sound more consonant than an octave. Although this has not been tested with a careful laboratory experiment, the clear sense from hearing clarinet tones is that the intervals that are the most consonant (the octave, for example) are the same as with other instruments. Once again, this must be because of the effect of prior exposure to other instruments and human speech, which do contain even-numbered harmonics. One can only speculate about how the various intervals would sound if our only auditory exposure was to clarinet-like timbres.

One is compelled to conclude that there is much more to the perception of harmony, and the schematic expectations that are such a salient part of it, than is determined by the structure of sound and the mechanical properties of the ear. But it is a mistake to discount these factors completely. Since the time of Helmholtz (1885/1954), the notion that relationships in harmony can be explained by aspects of the structure of sound and processes in early audition has been a leading hypothesis and remains so today (see Parncutt, 1989). It would be surprising if the conventions that have been established were chosen arbitrarily. Factors such as spectral overlap and the consonance and dissonance it engenders may have played a role in the early establishment of the conventions that today drive our perception. One can imagine musicians in ancient and medieval times experimenting with combinations of tones and finding certain combinations to be more consonant than others, based entirely on spectral overlap. The ensuing choices then became convention, and today we are so inundated with music that adheres to those conventions that our internalization of those conventions can compensate to some extent when spectral overlap fails.

How Expectations May Be Learned: Neural Net Modeling

The above evidence forces us to conclude that tonal expectations must be either innate or learned. There are obvious problems with the innateness hypothesis. First, although it is axiomatic that we must be innately endowed with a mechanism capable of supporting schematic expectations, it is difficult to think of the evolutionary pressures that would have given rise to the specific schematic expectations that occur in the perception of Western tonal music by Western listeners.

Cross-cultural studies would, of course, shed some light on how different or similar schematic expectations are around the world.[16] It is important to note, however, that universality does not imply innateness (see Bharucha & Olney, 1989). Studies with infants can only be suggestive because perceptual learning can occur very early, possibly even in utero.

Perhaps the most compelling argument against the innateness hypothesis is the existence of a plausible learning hypothesis. If schematic expectancies can be easily learned by mere exposure, there is no reason for them to have been wired innately. Much of the work in music cognition (as well as work on artificially intelligent musical systems) consists of postulating rules. This approach is most pronounced in the case of grammars. This work is invaluable because the rules serve as crystallizations of the constraints on the system. In particular, they characterize the knowledge that a musical schema must exhibit. Psychologists are often interested in also understanding the mechanisms that underlie the encoding of this knowledge. From this point of view, it is interesting to ask how the properties that are captured by rules are learned by passive exposure.

Recent advances in our understanding of how brainlike systems might learn provide an opportunity to explore how musical schemas might be learned (Rumelhart & McClelland, 1986). Models of brainlike systems are broadly referred to as neural nets, connectionist models, or parallel distributed models. These are, in turn, a special subset of a larger class of models that may simply be called associative networks.

In associative models, objects or features in the world are represented by units (or nodes). Representational units are linked together to form a network. The network as a whole represents the complex array of relationships between the objects represented by the units. The state of activation of a unit represents the degree to which the represented object is attended (suggested, implied, or expected).

If a network contains units that represent musical objects such as tones and chords, the connections between these units collectively encode the relationships that hold between these tones and chords as they are typically used. The state of activation of a unit represents the degree to which that tone or chord is currently being attended. A unit may be activated directly by a stimulus in the environment, such as a tone or chord, or indirectly through an associative network of such units. When a unit is indirectly activated, that tone or chord is suggested, implied, or expected.[17]

There are many arguments in favor of understanding music cognition in terms of associative networks. First, they easily accommodate ambiguities, with which music is rife. Many units can be active simultaneously to varying degrees, representing varying degrees to which different tones, chords, or keys are suggested. Second, they fill in missing pieces that are parts of typical patterns. For example, if the root and fifth of a triad are played, a network will suggest the major and minor third in proportion to their association with the prior context. Or if an F major chord is followed by a G major chord, the representational unit for C major will be strongly activated, representing a strong expectation for C major.

There are at least two additional arguments in favor of neural net models specifically. First, they are based on fundamental principles of neural organization in the brain. Second, and most important, they can learn.

Learning in a neural net model occurs by changing the strength of the connections between representational units. Consider a neural net that receives its input via units that represent octave-equivalent pitch classes (shown at the bottom of Figure 9.6a). Each such unit is said to be tuned to a particular pitch class. This is a valid assumption given that there are neurons in the auditory system that are tuned to specific frequency ranges and are laid out spatially according to frequency.[18]

Now suppose that this set of units is connected to another set of units that initially do not have any specialization (labeled "unspecialized" in Figure 9.6a). Each pitch class unit is connected to each unspecialized unit (the connections are not shown in the figure), such that the strengths of the connections are initially random. When a combination of tones is sounded, say an F-sharp major chord, the units representing the tones F-sharp, A-sharp, and C-sharp are activated (as shown by filled circles in Figure 9.6b). They in turn activate the unspecialized units. Since the connection strengths are initially random, the activations of the activation levels of the unspecialized units will be random.

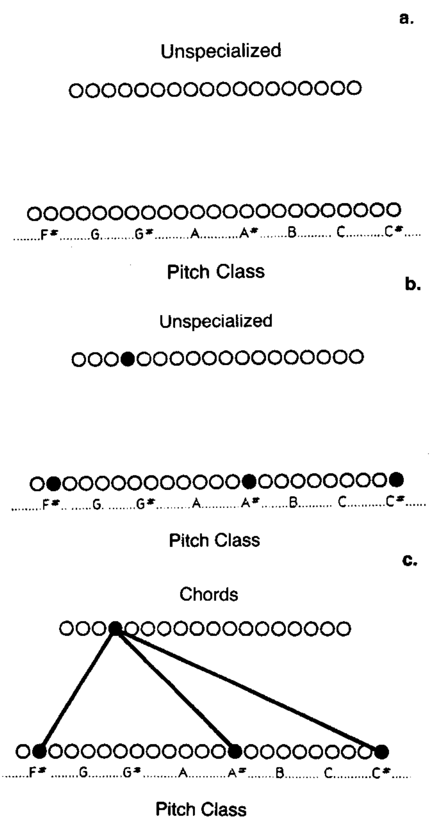

FIGURE 9.6. A neural network learning to represent typical combinations of tones. (a) Pitch class units in the bottom layer are connected to unspecialized units in the top layer via connections with random strengths. (b) In response to a combination of tones (F-sharp, A-sharp, C-sharp), the corresponding pitch class units are activated, and in turn activate the specialized units, one of which wins over the others. (c) The connections between the activated pitch class units and the winning unspecialized unit are strengthened, producing a unit that is now specialized for that combination of pitch classes.

A principle found in many parts of the brain is that units within a layer inhibit each other in proportion to their activation. The consequence of this is that only the most highly activated unspecialized unit remains active, and the others are shut down. So in response to the F-sharp major chord, some arbitrary unspecialized unit (the filled circle in Figure 9.6b) is going to win over the others in that set.

At this point, learning occurs by changing the strengths of the connections between the pitch class units and the winning unspecialized unit. The strengths are changed according to Hebb's (1949) rule: A connection becomes stronger in proportion to the product of the activations of the two units it connects. Since the F-sharp, A-sharp, and C-sharp pitch class units are the only pitch class units activated, their connections to the winning unit become strengthened (indicated by bold lines in Figure 9.6c) and the other connections to the winning unit become weakened. This winning unit has now become specialized to respond to F-sharp major chords. It will be even more likely to win the next time an F-sharp major chord occurs and will be even less likely to win in response to some other chord.

In this way, specialized representational units can form for all combinations of musical events that occur with great regularity. This process can be continued to yet another set of units to which the newly formed chord units are connected. This third set can become specialized to respond to typical combinations of chords, such as tonic, dominant, and subdominant.

Consider an idealized end-product of this learning scheme. It consists of pitch class units connected to units representing their parent chords, and chord units connected to units representing their parent keys, as shown in Figure 9.7 (see Bharucha, 1987a, 1987b, for details). The minor chord units are shown below the pitch class units for convenience only. Furthermore, the chords are shown in a systematic order for illustrative convenience only. What's important is the pattern of connectivity between units, not the layout of the units themselves.

After learning, the network can be tested to see if it has any interesting properties. When presented with the tones C, E, and G, units representing the chords that contain any of these tones are activated. These chord units in turn activate all units representing tones that are contained in these chords and keys that contain these chords. The activated key units in turn activate all units representing chords that are contained in these keys, and so on. The activation process eventually dissipates, and the network reaches an equilibrium state.

At this point, the activations of the chord and key units represent the degree to which chords and keys are suggested or expected. In response to three tones comprising a major triad, the expectations generated by the network are exactly in accord with the circle of fifths: chords and keys closest around the circle are most highly expected. In response to two major chords in succession whose roots are a whole tone apart (e.g., F major followed by G major), the network most strongly activates the key (C major) of which these chords are subdominant and

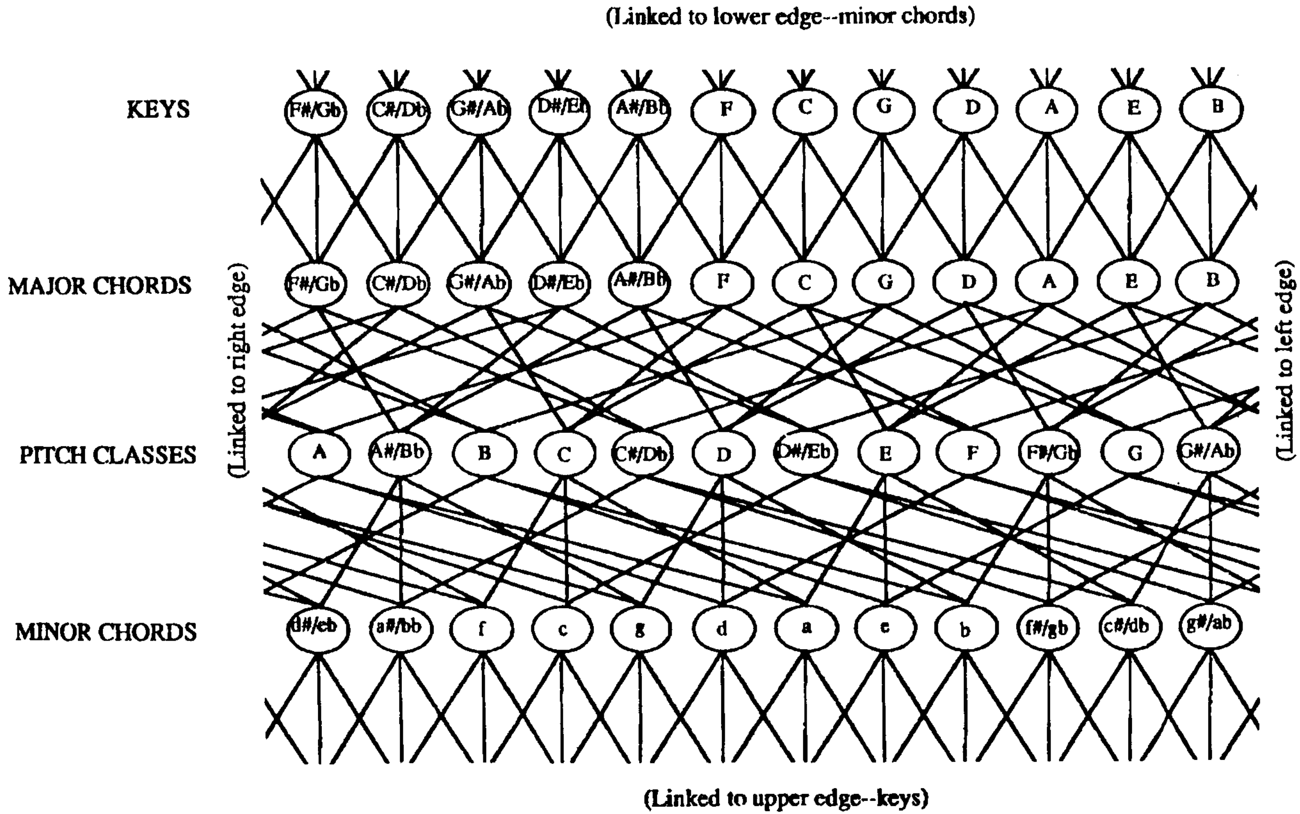

FIGURE 9.7. A network resulting from learning by exposure to typical combinations of tones as found in Western music. (Adapted from Bharucha, 1987a. Copyright by the Cognitive Science Society.)

dominant, and generates a strong expectation for the tonic chord. Finally, in response to a key-establishing context such as a major chord or a final cadence, the network generates expectations for the diatonic tones in that key. To this extent, the network can be seen to generate tonal hierarchies.

The network exhibits the patterns of expectation found in experiments on chord priming. The indirect activation of a chord unit represents an anticipatory processing of the represented object, resulting in faster and more accurate judgments about that chord, as observed in priming.

These complex behaviors of the network cannot be seen immediately by simply examining a diagram of the network such as Figure 9.7. They emerge from computer simulations of the activation process operating on the network. The network thus has emergent properties—that is, properties that are not directly evident from the initial assumptions but emerge from the way in which these assumptions interact. Emergent properties that are in accord with experimental data and music theory, as they are in this case, constitute strong support for a theory. Most theories that consist of rules (such as Lerdahl and Jackendoff's theory) postulate those rules explicitly and do not provide an account of how those rules are acquired. In a neural net, in contrast, the system behaves in accord with rules that weren't postulated explicitly but instead emerge from the complex interactions of more elementary processes that are themselves learned.

Neural nets thus offer the potential to understand how tonal schemata may be learned by passive perceptual learning, thereby giving rise to the schematic expectations that form such an integral part of our musical experience.

NOTES

1. These include (but are by no means limited to) Zarlino (1558/1968), Fux (1725/1943), Schenker (1935/1979), Schoenberg (1954/1969), Meyer (1956), Lerdahl and Jackendoff (1983), Narmour (1990), and Butler and Brown (Chapter 8 in this volume).

2. As opposed to rhythmic or melodic expectations.

3. In keeping with the goals of this book, I shall attempt to avoid detail and excursions in favor of elucidating a clear line of reasoning. The reader is directed to other sources for details, for additional work, and for alternative explanations that bear on the question of harmonic expectation. These include Deutsch and Feroe (1981), Sloboda (1985), Dowling and Harwood (1986), Handel (1989), and, most relevant to some of the material covered in this chapter, Krumhansl (1990). For details about the work of the author, see Bharucha (1987, 1991; Bharucha & Todd, 1989).

4. It should be noted that there is some debate about whether "emotion" is the appropriate term to describe the often subtle and ineffable experiences induced by the fulfillment and violation of tonal expectations. Whether one is willing to call these experiences emotions, however, they are clearly affective.

5. There is no known reason to suppose that one cannot have more than one musical culture.

6. Many cognitive psychologists would consider a schema to be a memory system,

too. For the sake of simplicity, however, we shall restrict the term "memory" to memory for pieces of music.

It may be possible for the schematic expectations to emerge from the memory instead of requiring a completely distinct structure. One way in which this might occur has been explored by Bharucha and Todd (1989). Other methods might generate the schema by an average or composite (e.g., Metcalfe, 1982) of all individual memory traces. Whether the schema and the memory are part of the same physical structure or separate structures is still an open research question. In this chapter it will be assumed that they are separate.

7. This conclusion must await further research, however, since the empirical determination of this issue is very tricky. When a piece of music is recalled from memory in order to perform or imagine it, the schema is activated just as it would be by the sound itself, thereby generating the event hierarchy. It therefore seems redundant to store the elaboration in memory, since it is generated automatically every time the piece is recalled. This is what makes an empirical answer to the question so difficult.

8. This is based on the probability with which chords follow each other in the musical genre to which the subjects have been most extensively exposed (Western tonal music for the subjects in these experiments). C major is most often followed by G major or F major and least often by F-sharp major. Hence, for these subjects, G and F are objectively more highly expected than is F-sharp, following C.

9. The figure shows root position triads in close position as an example only. Although the original experiments employed "Shepard" tones (Shepard, 1964), which render tones more ambiguous with respect to their particular octave, subsequent experiments (Stoeckig, 1990) have shown that the results from this priming experiment are quite robust for different inversions, positions, and voice leading. Thus the effect seems to truly capture the perceptual properties of the triads as wholes.

10. In the case of the linear spatial representation of pitch height this is true! Such a *tonotopic* representation of pitch height is found not just in the inner ear, but even in the cerebral cortex.

11. The information contained in a circle can be encoded by a computer in many possible ways, including a string of symbols specifying the mathematical equation of the circle. Similarly, the brain might encode this information in terms of a particular arrangement of neurons, even though this arrangement may not itself be circular. Cognitive psychologists explore mental representations at different levels of information abstraction, spatial representations being some of the most abstract and neural representations some of the least abstract. The advantage of abstract representations is that the essence of the information can be captured without worrying about the representational medium. The advantage of less abstract representations is that the causal mechanisms involved in encoding the information can be addressed.

12. The strength of a correlation of two profiles is given by a single number, the correlation coefficient, r, that ranges between -1 and 1. A correlation coefficient of zero signifies no correlation.

13. There is nothing mystical about a four-dimensional space. Mathematically, to say that more than three dimensions are required to capture all the relationships between major and minor keys simply means that more than three factors are needed for a complete description of these relationships. The difficulty of depicting the four-dimensional key space stems only from an attempt to illustrate the space in terms of a model that can be visualized.

14. Parncutt's (1989) theory of the psychoacoustic foundations of harmony predicts this outcome.

15. This is not the same as being able to imagine music after becoming completely deaf, which only requires auditory imagery. The familiar account of Beethoven continuing to compose while deaf doesn't in and of itself discount the spectral overlap hypothesis. If sounds are stored in memory in terms of their spectra, then, while imagining music, the imagined spectra would overlap or not, just as would real spectra.

16. See Bharucha (1987b) for a report of a preliminary priming experiment using Indian music and Indian subjects.

17. See Deutsch (1969) for an early model of this kind. Unlike more recent models, it couldn't learn, since the mechanisms by which neural nets might learn were not then well understood.

18. This is called a tonotopic mapping, and can be found all the way from the inner ear to the auditory cortex.

REFERENCES

Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York: Freeman.

Bharucha, J. J. (1984). Event hierarchies, tonal hierarchies, and assimilation: A reply to Deutsch and Dowling. *Journal of Experimental Psychology: General, 113*, 421–25.

Bharucha, J. J. (1987a). MUSACT: A connectionist model of musical harmony. In *Proceedings of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Bharucha, J. J. (1987b). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception* 5, 1–30.

Bharucha, J. J. (1991). Pitch, harmony, and neural nets: A psychological perspective. In P. Todd & D. G. Loy (Eds.), *Music and connectionism*. Cambridge: MIT Press.

Bharucha, J. J., & Olney, K. L. (1989). Tonal cognition, artificial intelligence and neural nets. *Contemporary Music Review*, 4, 341–56.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 403–10.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception and Psychophysics, 41*, 519–24.

Bharucha, J. J., & Stoeckig, K. (1989). *Chord priming: The automaticity of schematic expectancies*. Paper presented at the meeting of the Psychonomic Society, Atlanta, GA.

Bharucha, J. J., & Todd, P. (1989). Modeling the perception of tonal structure with neural nets. *Computer Music Journal, 13*(4), 44–53.

Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology, 1*, 12–29.

Carlsen, J. C., Divenyi, P. L., & Taylor, J. A. (1970). A preliminary study of perceptual expectancy in melodic configurations. *Council of Research in Music Education*, 22, 4–12.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology; General, 113*, 394–412.

Deutsch, D. (1969). Music recognition. *Psychological Review, 76*, 300–7.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88*, 503–22.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. San Diego: Academic Press.

Fux, J. J. (1943). *Steps to Parnassus* (A. Mann, Ed. and Trans.). New York: Norton. (Original work published 1725)

Geary, J. M. (1980). Consonance and dissonance of pairs of inharmonic sounds. *Journal of the Acoustical Society of America, 67*, 1785–89.

Handel, S. (1989). *Listening*. Cambridge: MIT Press.

Hebb, D. O. (1949). *The organization of behavior*, New York: Wiley.

Helmholtz, H. (1954). *On the sensation of tone*. New York: Dover. (Original work published 1885)

Jairazbhoy, N. (1971). *The rags of North Indian music: Their structure and evolution*. London: Faber & Faber.

Jones, M. R. (1981a). Music as a stimulus for psychological motion. Part I. Some determinates of expectancies. *Psychomusicology*, 2, 34–51.

Jones, M. R. (1981b). Music as a stimulus for psychological motion. Part II. An expectancy model. *Psychomusicology*, 2, 1–13.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11*, 346–74.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford: Oxford University Press.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334–68.

Krumhansl, C. L., & Schmuckler, M. A. (1986). The *Petroushka* chord: A perceptual investigation. *Music Perception*, 4, 153–84.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 579–94.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge: MIT Press.

Mandler, G. (1984). *Mind and body*. New York: Norton.

Metcalfe Eich, J. (1982). A composite holographic associative recall model. *Psychological Review, 89*, 627–61.

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

Narmour, E. (1977). *Beyond Schenkerism: The need for alternatives in music analysis*. Chicago: University of Chicago Press.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures*. Chicago: University of Chicago Press.

Parncutt, R. (1989). *Harmony: A psychoacoustic approach*. Berlin: Springer-Verlag.

Pierce, J. R. (1969). Attaining consonance in arbitrary scales. *Journal of the Acoustical Society of America, 40*, 249.

Pierce, J. R. (1983). *The science of musical sound*. Scientific American Books.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vols. 1 & 2). Cambridge: MIT Press.

Schenker, H. (1979). *Free composition* (E. Oster, Ed. and Trans.). New York: Longman. (Original work published 1935)

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception*, 7, 109–50.

Schoenberg, A. (1969). *Structural functions of harmony* (L. Stein, Ed.). New York: Norton. (Original work published 1954)

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 36, 2346–53.

Sloboda, J. A. (1985). *The music mind.* Oxford: Oxford University Press.

Stoeckig, K. (1990). *The effects of octave transformations: Evidence from harmonic priming concerning the functional equivalence of chords and their inversions.* Unpublished doctoral dissertation, Department of Psychology, Dartmouth College, Hanover, NH.

Tekman, H., & Bharucha, J. J. (1991). *Dynamics of the development of harmonic priming: From tones to keys and back.* Paper presented at the meeting of the Eastern Psychological Association, New York.

Tramo, M. J., Bharucha, J. J., & Musiek, F. E. (1990). Music perception and cognition following bilateral lesions of auditory cortex. *Journal of Cognitive Neuroscience, 2,* 195–212.

Zarlino, G. (1968). *The art of counterpoint* (G. A. Marco & C. V. Palisca, Trans.). New Haven, CT: Yale University Press. (Original work published 1558)