Tonal Hierarchies and Rare Intervals in Music Cognition

CAROL L. KRUMHANSL
Cornell University

Four issues raised by Butler’s (1989) commentary are addressed. The first issue is the possibility that the results of perceptual studies of tonal hierarchies can be attributed to task-specific strategies developed in response to particular stimuli. Such strategies cannot account for the convergence across experiments employing varied tasks and stimulus materials. The second issue is the correspondence between statistical summaries of music and perceptual data. The correspondence is shown to be quite general and to have implications for the acquisition of tonal knowledge. The third issue is the process listeners use to identify the tonal center. Pattern-matching to tonal hierarchies is shown to be a plausible process contributing to key-finding, whereas a tritone rule has limited applicability. The final issue is the effect of temporal order on pitch perception. Principled temporal-order effects are found in many psychological experiments, but not in those focusing on the tritone relation.

Butler’s (1989) commentary attempts a critical evaluation of two theories, seen in opposition, of cognitive processes in musical pitch perception. The first is called the “tonal hierarchy theory,” although the exact referent of this term is unclear. As discussed below, it appears to be used to refer to experimental results showing perceptual differentiation of tones in a tonal hierarchy. It is important to emphasize from the outset that the tonal hierarchy is just one component of experienced listeners’ abstract knowledge of relations among tones, chords, and keys demonstrated by experimental studies. Psychological data also describe the perceived relations between tones, between chords, and between keys; the hierarchical ordering

Requests for reprints may be addressed to Carol L. Krumhansl, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853.
of harmonic stability; and the harmonic functions of chords in keys. A central assumption underlying research in this tradition is that, through experience, listeners have internalized a richly structured system of knowledge about stylistic regularities in tonal-harmonic music. During music listening, the sounded events are interpreted in terms of this knowledge, which in turn influences musical memory, the formation of larger structural units, and expectations for future events. The "theory of intervallic rivalry" is proposed as an alternative approach. Its primary goal, considerably more limited, is to describe how a listener arrives at a sense of key from a sequence of sounded tones. The central claim is that the "rare" intervals (Browne, 1981) in the diatonic set, tritones and minor seconds, provide the clearest evidence concerning tonality and, thus, constitute the most important cues for key-finding.

To establish a basis for discussion, it might be well to begin by noting certain points in Butler's description of the "tonal hierarchy theory." He focuses on two representations: the conical representation of intertone distances and the graphical representation of the probe-tone judgments. Of the former, Butler (1989, p. 222) says, "Krumhansl (1979, 1986 [1987b]; Krumhansl & Shepard, 1979 [incorrect reference]) proposed an alternative geometrical model [to Shepard's 1982a, b model], in the form of a cone, to describe the hierarchical nature of perceived stability in the various tones of the diatonic [chromatic] set. The tonal hierarchy theory depicts the tonic (positioned at the vertex of the cone) as the most stable and consonant member of the major diatonic set; next-most stable are the dominant . . . and the mediant . . . pictured near the vertex of the cone. The remaining tones in the diatonic set are considered less stable and positioned farther from the vertex, and nonmembers of the set are deemed least stable, and positioned farthest-removed from the cone's vertex."

This description reflects a basic misconception concerning the nature of the representation and, more generally, the fundamental role of experimental data in psychological research. This conical representation is not a theoretical model, but instead a summary of psychological data. It derives from an experiment (Krumhansl, 1979) in which all possible pairs of chromatic scale tones in an octave range were presented following major scale and major triad contexts. Listeners were instructed to rate how similar the first tone is to the second, and the resulting matrix of rating judgments was analyzed by using multidimensional scaling. This method of analysis produces a spatial configuration of points such that interpoint distance is inversely related to psychological similarity. In this case, the scaling analysis produced a three-dimensional configuration with the tonic, dominant, and mediant located close together next to the vertex of a cone, the other scale tones farther from the vertex, and the nonscale tones farthest from the vertex. The tones were ordered around the cone as on the chroma circle.

None of these characteristics, including the conical shape itself, was de-
terminated from theoretical considerations; the spatial configuration was unconstrained except by the similarity judgments themselves. The significance of the multidimensional scaling analysis was to show that the perceived similarity judgments could be accounted for in terms of two underlying factors: distance on the chroma circle (represented as distance around the cone), and a tonal hierarchy corresponding to the music-theoretic notion of stability (represented as distance from the vertex). The latter should be distinguished from the notion of consonance for reasons detailed in Krumhansl (1990). Thus, the scaling analysis of the data served the purpose of uncovering factors governing the perceived similarity between tones presented in a tonal context.

The tonal hierarchy dimension of the scaling solution was first evident in the results of experiments using the probe-tone method (Krumhansl & Shepard, 1979). Of this method, Butler (1989, p. 223) says, "Most experimental studies of the tonal hierarchy have asked listeners to rank [rate] how well various tones or chords [not applicable] are judged to 'fit with' short contextual patterns such as scales and major chords (e.g., Krumhansl, 1979 [the correct reference would be Krumhansl and Shepard, 1979, although the instructions were somewhat different], Krumhansl & Kessler, 1982). These studies have shown that the tonal hierarchy, as manifested both in probe-tone ratings and in statistical distributions of tones in compositions thus analyzed, tend to take one of the shapes shown in [his] Figure 2. The tonic is defined as the most stable member of the set; next in rank order are the dominant, the mediant [this order is reversed for minor contexts], the remaining set members, and then nonmembers of the set."

Here, again, Butler appears to be confusing data with theory. The values reproduced in his Figure 2 (and shown as the dashed lines of Figure 1 here) are the data from the experiment rather than a theoretical model as implied. The tones rated highly in the probe-tone task are the same as those located close together in the scaling analysis of tone-similarity judgments. This convergence at the level of data is important for establishing the tonal hierarchy as a basic psychological factor; the coincidence is not imposed theoretically. The correspondence between these data and statistical distributions of tones in compositions, and its significance, will be discussed later. With this background, we can turn now to four general issues raised in Butler’s (1989) commentary. In certain instances, specific methodological and technical points will also be considered for purposes of clarification.

Abstract Long-Term Representations of Pitch Structure vs. Short-Term Effects of the Perceptually Immediate Context

One issue raised by Butler (1989; see also Butler 1988b) is whether experiments such as these which demonstrate hierarchical differentiation of
tones reflect long-term knowledge of pitch structure or simply special strategies developed in response to the particular task and stimulus materials. He is concerned primarily with the probe-tone rating task, in which a context is followed on different trials by all 12 tones of the chromatic scale. To be concrete, Butler claims that the "typical" contextual sequence shown in his Figure 7 (which, to my knowledge, has never been employed in any published experiment) might induce a strategy to rate the tonic most highly because it is sounded twice, once at the beginning of the scalar sequence and once at the end, and with longer durations than the remaining diatonic tones. Even if these factors were influential, it would not account for the results he reproduces in his Figure 2. These results show differentiation among the remaining (nontonic) scale tones (e.g., higher ratings of the dominant than the subdominant), whereas his account would predict these to be equal.

Arguments of this kind that probe-tone ratings simply reflect influences of such factors as tone repetition, duration, and serial position are unsupported. The contexts actually used in the probe-tone studies (e.g., Krumhansl & Shepard, 1979; Krumhansl & Kessler, 1982; Cuddy & Badertscher, 1987) are varied: incomplete ascending and descending scales (without the final tonic), complete scales (with the final tonic), tonic triads (blocked and arpeggiated), and chord cadences. Despite these variations, the agreement between results from the different conditions evidences a robust and reliable pattern that cannot be attributed to the particulars of the contexts. In fact, the different contexts were used in the experimental designs for the purpose of testing for agreement in the rating data despite variations in the particular contexts used. Krumhansl and Kessler (1982), for example, reported a high correlation between the probe tone ratings for four different major key contexts (.896) and four different minor key contexts (.910). These values were considerably higher than would be predicted from correlating the tone distributions in the different contexts (the correlations average .752). This shows that the consistency of the data cannot be attributed to stimulus factors alone. Had Butler done this appropriate analysis, rather than summing tone distributions (his Figure 8), he would have been unable to conclude "this similarity [between probe tone ratings and tone distributions] reflects a stimulus artifact in the design of Krumhansl and Kessler's experiment, rather than a reflection of mental representations of pitch relations" (pp. 230–231).

Butler also disregards data arising from experiments that use different tasks. For example, tonal hierarchies are evident also in memory confusions (e.g., Cuddy, Cohen, & Miller, 1979; Krumhansl, 1979), reaction times to judge key membership (Janata & Reisberg, 1988), judgments of what constitutes a good or complete phrase (Palmer & Krumhansl, 1987a,b), and expectations for melodic continuations (Schmuckler, 1988). The last three
studies, as well as the probe-tone experiment of Thompson (1986), employ segments of musical compositions, demonstrating that these evoke similar abstract tonal hierarchies. In short, Butler’s conclusion that “serious problems” of a methodological nature attend the “atonal hierarchy theory” are readily disconfirmed by considering the extent to which short-term effects of the perceptually immediate context fail to predict the experimental data at hand, and the range of tasks and contexts generating a similar pattern of hierarchical differentiation of tones.

Butler’s (1989, p. 233; see also Butler, 1988b) account of the second probe-tone study of Krumhansl and Kessler (1982) reveals a number of misconceptions about how the data in that study were analyzed and interpreted. These aspects are important for understanding the approach taken to evaluate the extent to which the results reflect local effects of individual chords. More generally, they illustrate how the probe-tone methodology can be used as a tool for uncovering different levels of musical organization. The objective of the study was to trace how the sense of key develops and changes over time as sounded information is evaluated and integrated by the listener. For this purpose, we obtained probe-tone ratings after each successive chord in various sequences, some of which suggested modulations to more or less closely related keys.

The ratings at each position were correlated with the tonal hierarchies obtained separately for unambiguous key-defining contexts; these values (rather than “rankings of well-fittedness of probe tones”) were taken as the measure of key strengths. These were compared with analogous correlations by using probe-tone ratings for isolated chords also obtained separately. (This is how we used the data Butler, 1989, p. 229, footnote 10, claims “were dropped from the analysis without explanation.”) These two sets of values are plotted in Figure 8 of Krumhansl and Kessler (1982), where it is explained: “The difference between the graphs . . . is a measure of the extent to which the listener perceives the prevailing tonal organization as opposed to the tonal organization of the last heard chord” (p. 356).

Treating the data in this way enabled us to separate effects of tonicization at one level from a sense of the prevailing key at another level. This was related to Schenker’s (1906/1954, 1935/1979) theory of hierarchic levels of tonal organization. Thus, it seems unfair to conclude that the probe-tone method is unable to separate surface from more abstract levels of organization (Butler, 1989, p. 223) or to account for the process of tonicization (Butler, 1989, pp. 239–240).

This and other studies show that the perceptually immediate context has an effect on the musical experience. It would be strange, indeed, if the sounded events did not have special salience. What the experimental studies reviewed to this point demonstrate is that, in addition, tonal contexts (even simple, schematic ones) evoke listeners’ knowledge about how musical en-
tities, such as tones, chords, and keys, are generally used in the style. These abstract pitch relations provide a background that facilitates understanding the organization of a particular piece: the roles of sounded events in larger melodic, harmonic, and rhythmic patterns; the building and release of tension; modulations to more or less closely related keys; and so on. As described by Bharucha (1984b), each event has a place not only in the abstract tonal and harmonic hierarchies (referring to classes of tone and chord instances), but also a place in the event hierarchy of the particular piece (referring to the string of musical events that constitutes a piece of music).

Moreover, a number of studies suggest that surface salience (resulting from such factors as tone repetition, duration, overall pitch height, serial position, metrical stress) may be important for enabling the listener to discover principles of pitch organization in novel musical styles. For example, Castellano, Bharucha, and Krumhansl (1984) found style-appropriate tonal hierarchies for North Indian music with inexperienced listeners, which was attributed to the explicit emphasis in the music given to structurally significant tones (as summarized in tone durations). Similar findings were obtained in Kessler, Hansen, and Shepard’s (1984) study of Balinese music. Effects of surface salience also were found in studies of octatonic (Krumhansl & Schmuckler, 1986b) and 12-tone serial writing (Krumhansl, Sandell, & Sargeant, 1987). In all these cases (with the last being a possible exception), surface salience appears to initiate the abstraction of style-appropriate pitch organizations in music outside the listener’s prior experience. With increased exposure, a process of this sort is assumed to establish long-term representations of compositional principles that can be evoked with simple, schematic contexts.

**Statistical Distribution of Musical Elements and Psychological Representations of Pitch Structure**

The next issue to be discussed is the utility of statistical surveys of musical compositions in connection with psychological research. The correspondence between the profile of total durations of the 12 chromatic scale tones in Schubert, op. 91 no. 1 (Hughes, 1977) and psychological measures of the tonal hierarchy is striking; the correlation between the tone durations and the tonal hierarchy of G major is .97. Hughes (1977, p. 157) states: “Considered as a whole, the results of op. 94 no. 1 indicate a rather simple tonal complexity much closer to the key of G major than to C major as presented by Schubert.” Butler’s (1989, p. 225) confusion on this point (“Hughes (1977) stated that his sums of note durations indicated that the music was oriented “toward” G major although the key of the composition was C major, and indeed Schubert seemed to believe the piece was in C ma-
The analysis of tone distributions in this particular piece raised the question of the generality of the correspondence between distributions of elements in music and psychological data on perceived tonal and harmonic structures. As summarized in Krumhansl (1987b; and more extensively in Krumhansl, 1990), the correspondences are generally strong. Figure 1 shows the tone frequencies in a variety of melodies (a sample of nearly 25,000 tones) from studies by Youngblood (1958) and Knopoff and Hutchinson (1983) and the tonal hierarchies (Krumhansl & Kessler, 1982) of the appropriate key. (In the figure, the values are plotted with respect to a reference tonic of $C$.) Similar correspondences obtain between frequencies of successive pairs of tones and judgments of perceived relatedness between tones, frequencies of chords and perceived harmonic hierarchies, and frequencies of harmonic progressions and judgments of perceived relatedness between chords. These analyses show a strong, general correspondence between distributions of elements in musical compositions and psychological data. Naturally, the relationships would be weaker if, for example, the statistical sample were based on a composition involving complex modulations, on short segments for which the distributions would be less stable, or on a segment of a composition intended to set up conflicting or ambiguous tonal cues.

What significance do these correspondences have for music perception? Butler (1989, p. 222) states: "The tonal hierarchy theory has been invoked to account for . . . statistical distributions of tones within actual musical compositions (Krumhansl, 1986 [Krumhansl, 1987b])." In fact, just the opposite was proposed: "These correspondences suggest that knowledge of tonal practice evident in listeners' judgments is abstracted from their musical experience, and that through this experience they have internalized the regularities of the musical style" (Krumhansl, 1987b, p. 22). This hypothesis is congruent with a wide variety of results showing that humans (as well as other organisms) are highly sensitive to information about frequency of occurrence. Thus, the primary significance of the observed correspondence between statistics of music and psychological data in these cases is to suggest a mechanism through which principles of musical organization are learned.
Statistical summaries also can guide psychological experimentation by helping to identify typical (as opposed to infrequent) pitch patterns for study. For example, a glance at Figure 1 (as well as the distributions presented by Butler, 1989) shows that the diatonic tritone in major keys (between the subdominant and the leading tone) is much less frequently represented in musical compositions than most all other diatonic intervals. This raises the question of what process might enable listeners to learn the spe-
Tonal Hierarchies and Rare Intervals  

317

cial significance of the tritone as a cue for the key of a piece of music. Finally, statistical summaries can help relate musical and psychological constructs to one another. The literature contains a number of successful applications of this approach. Lindblom and Sundberg (1970) used the rate of tonal events to substantiate a generative theory of constituent structure. Hutchinson and Knopoff (1979) related the predominance of chord voicings with larger intervals in the bass than the treble to increases in critical band width in lower registers. Recently, Vos and Troost (1989) found ascending intervals tended to be larger than descending intervals, consistent with Meyer’s (1973) gap-fill melodic process and the idea that the large ascending interval induces a feeling of tension or arousal that is reduced when the gap is subsequently filled in by smaller descending intervals. Although surveys of features found in musical compositions can only provide information about the way elements are generally used in a musical style, they can serve to substantiate the operation of certain psychological processes.

The Process of Key-Finding

A variety of psychological measures indicate that listeners experienced with tonal-harmonic music tend to interpret the sounded tones and chords in terms of their functions in the prevailing key. This raises the question of how listeners initially arrive at a sense of key. Intuitively, it seems that certain cues may be important: chord cadences, tonic-dominant-median relations, the diatonic set, and the “rare” intervals of tritones and minor seconds. In evaluating these possibilities, it seems important to bear in mind that they are not mutually exclusive. Listeners may use a number of cues in combination, with the weighting possibly dependent on the musical passage in question.

The correspondence described in the last section between tonal distributions and tonal hierarchies suggests that listeners may develop a sense of key, in part, by matching the sounded events in a particular passage to the learned tonal hierarchies. A computer algorithm (Krumhansl & Schmucker, 1986a; Krumhansl, 1987b, 1990) was developed to test the efficacy of a key-finding process of this sort. This simple algorithm is based on matching the distribution of tone durations in the input segment to the tonal hierarchies of the 24 major and minor keys. This is done by correlation, which matches the patterns as wholes rather than by piecemeal comparison of relative rankings as was done by Butler (1989, pp. 226–228). The key with the highest correlation is said to be the key found by the algorithm. Essentially, this algorithm simulates a process of pattern-matching between the durations of the tones in a musical excerpt and the psychological tonal hierarchies. In tests of the algorithm, the tone durations were taken from the notated score as a simple measure of the tonal content
of the passage. This measure ignores all other factors such as temporal order, harmonic implications, and metrical and rhythmic stress. How these factors might be incorporated in such an algorithm to sharpen the key-finding process is discussed in Krumhansl (1990). Various theoretical constructs, such as prominence profiles (Lindblom & Sundberg, 1970) and reductions (Lerdahl & Jackendoff, 1983), also may prove useful in filling out a description of the key-finding process.

Even without taking these other factors into account, tests of the algorithm showed that matching to tonal hierarchies yields fairly accurate results—and this is true for very short musical excerpts. For purposes of later comparisons with a tritone rule (applicable only to major keys), I will focus on the applications of the algorithm to pieces in major keys. One test of the algorithm used just the first four sounded tones in the preludes from Bach’s Well-tempered Clavier, and the preludes of Shostakovich and Chopin. The algorithm determined the correct key for 79% of the major preludes using just the first four tones. In an application to the fugues of Bach and Shostakovich, the tones of the fugue subjects were added one by one. The correct major key was determined in 92% of the cases, with an average of 3.7 tones needed. A final application showed the algorithm successfully traced the shifting pattern of tonalities in Bach’s C minor prelude (Book II) in a manner that agreed well with judgments of key strengths by trained music theorists; in this case, the analysis was done on a measure-by-measure basis. Thus, tonal distributions in short excerpts contain a great deal of information about the key when matched to tonal hierarchies, contrary to Butler’s (1989, p. 226) claim that, based on statistical distributions of tones and tonal hierarchies, “a listener could identify the key of a piece of music with full certainty only after having heard and summed the durational weightings of all the tones in the composition—that is, after the performance was over.”

How effective might a tritone rule be for determining key? Butler and Brown (1984) reported an experiment in which listeners were 83–91% accurate in identifying the tonic after hearing three-tone sequences including the unique tritone of a major key (formed by the subdominant and leading tone). They conclude (p. 12) the result “provides persuasive evidence that the tritone serves as a primary indicator of tonal center,” ignoring the finding that the same listeners were 96–98% accurate in choosing a possible tonic when the tritone was absent. When Butler (1989, p. 234) says of this study that “agreement erodes when rare intervals are absent,” it is presumably simply because the three-tone sequences without tritones allow multiple key interpretations. To measure correctly the contribution of the tritone, what is needed is a design in which two possible alternative tones are presented, one a possible tonic and one an impossible tonic. If the percentage of correct tonic judgments is higher for tritone sequences, this would support Butler and Brown’s (1984) conclusion. Unfortunately, a design
such as this has not been employed. Other studies (e.g., Butler, 1982, 1988a) do not permit the contribution of the tritone to be evaluated, because suitable control conditions without tritones, essential for testing the theoretical claim, are lacking. (Incidentally, when Butler, 1989, p. 239, reports of the 1988a study “responses to patterns derived from tonal compositions approached unanimity for every pattern but one,” he is talking about a total of three patterns. In addition, the motivation for including three atonal selections in a study of key-finding is obscure.)

Let us suppose for sake of argument, however, that a tritone rule might be one mechanism listeners use to arrive at a sense of key. It has the appeal of providing potentially definitive information (assuming major mode and no chromatic alterations). One can then go on to ask how often the music would offer an opportunity to apply such a rule. In the initial four-tone segments of the Bach, Shostakovich, and Chopin major preludes the tritone appears just once—that is, in 2% of the cases. [That Cohen’s (1977) listeners were 83% accurate judging the key of a subset of the Bach major preludes after hearing just the first four events eliminates the tritone as necessary for key-finding.] In only 58% of the major fugues of Bach and Shostakovich do the two tones of the tritone appear anywhere (not necessarily successively) in the entire fugue subject. When they do appear it is after an average of 12.2 tones. Thus, opportunities to apply a tritone rule occur relatively late and infrequently in the initial segments of these pieces, even though “Everyday musical experience tells us that tonally encultured listeners can recognize the tonal center in an unfamiliar composition almost instantly” (Butler, 1989, pp. 237–238).

Other observations suggest the applicability of a tritone rule is limited:

1. It can rarely be applied to monophonic music [according to Butler, (1988a) tritones appear predominantly as harmonic intervals; the tritone constitutes only 0.3% of Youngblood’s (1958) sample of nearly 2,000 melodic intervals].
2. It cannot be applied unambiguously to pieces in minor keys (the minor scale in both harmonic and ascending melodic forms contains two tritones).
3. It cannot be applied to other church modes or scales in other cultures (e.g., 4 of the 10 thāts—that is, scales—on the circle of thāts in North Indian music do not have a unique tritone; see Krumhansl (1987a)].
4. The tritone is relatively difficult to encode, label, remember, and produce.
5. A process keying-off (as it were) from tritones would be subject to false alarms when chromatic alterations occur [the only tonal example offered by Butler (1989), Figure 16, is a case in point].
Finally, Butler's illustration (1989, footnote 13) of how a listener might use "important tonality cues" suggests a concise, testable model (one that might, for example, be embodied in a computer algorithm) may prove elusive.

In my view, the next most rare interval in the diatonic set, the minor second, has far greater potential as a cue to key-finding, although this has not been the focus of Butler's experimental efforts to date. The positions of minor seconds are important for defining major and minor keys, as well as church modes, and thàts in Indian music. More generally, the relative location of large and small intervals also seems to be an important feature of pentatonic scales and Indonesian modes. Finally, minor seconds are relatively easy to produce, label, and remember and appear relatively frequently in diatonic music [constituting 22% of Youngblood's (1958) sample of melodic intervals, the majority of which, 62%, are diatonic minor seconds—leading-tone and tonic, mediant and subdominant]. Thus, the minor second appears to be a much more promising beginning for a theory of key-finding than the tritone. Again, however, a process based exclusively on these intervals would be subject to errors when chromatic alterations occur.

What is more important to note, however, is that emphasizing minor seconds, and the leading-tone to tonic relationship in particular, is in no way incompatible with results of experiments in the tradition of the "tonal hierarchy theory." For example, Krumhansl (1979, p. 354) found in a C major context, of the 132 melodic intervals tested (all ordered pairs of tones in an octave range), the highest similarity rating went to the interval BC. Moreover, the key-finding algorithm based on pattern-matching to tonal hierarchies (described earlier) finds C to be the strongest candidate for tonic when the input segment contains just these two tones. In summary, the "theory of intervallic rivalry" may be placing disproportionate emphasis on the tritone relation. Although its presence might eliminate residual ambiguity as to the key, other cues may be just as, or more, effective. Moreover, the circumstances under which a tritone rule might be used appear to be restricted.

Pitch and Time

The final, general issue to be addressed is Butler's (1989) plea to consider the manner in which pitch relations are temporally arrayed in music and its consequences for music perception. This is a centrally important issue and one that invites further theoretical and empirical analysis. Unfortunately, the experimental evidence he offers on this point is inconclusive and contradictory. Brown and Butler (1981; also reported in Butler & Brown, 1984, as Experiment 1) presented all six possible orders of three-tone sets containing the tritone. This study found an advantage in key-identification accu-
racy when the subdominant was followed by the leading tone over the reverse temporal order. This result is attributed to the possibility that the subdominant to leading-tone sequence implies the subdominant to dominant harmonic progression. However, Butler (1982, also reported in Butler & Brown, 1984, as Experiment 2) found, using paired dyads, higher accuracy for simultaneously sounded tritones than for successively sounded tritones, suggesting that temporal order of these tones may not contribute importantly to key-finding after all. Moreover, comparing the two cases in which the tritone is presented successively, there is now an advantage for the leading-tone to subdominant order over the subdominant to leading-tone order, contrary to the earlier study. Thus, the results of studies focusing on tritones do not yield a clear pattern of temporal-order effects.

In contrast, consistent, interpretable, and principled temporal-order effects have been obtained in other experiments. They were a subject of interest to a number of early psychological investigators (e.g., Meyer, 1903; Bingham, 1910; Farnsworth 1925, 1926a,b). These studies found regular differences in ending preference judgments of isolated pairs of tones as a function of their temporal order. Meyer favored an explanation based on frequency ratios, whereas Bingham and Farnsworth favored a learning-based explanation in which the listener attempts to match the intervals to familiar tonal relationships.

When pairs of tones were presented following a key-defining context, Krumhansl (1979, p. 358) found: “The similarity ratings of pairs of tones were often asymmetric, that is, the similarity ratings depended on the order in which the tones were presented. . . . The asymmetries followed a regular pattern . . . higher average similarity ratings were given when the first tone was less closely related to the tonality established by the context and the second tone was more closely related to the tonality than when the tones appeared in the opposite temporal order.” This result was related to dynamic tendencies as described by music theorists: “The tendency of the less structurally stable musical tones to move toward the tonic and its closely related tones in time is unanimously regarded by music theorists . . . as important in shaping the flow of music in time” (p. 363). It was also noted that such temporal-order effects cannot be captured in geometric representations in which distances are necessarily symmetric. Various approaches to this representational problem are considered elsewhere (Krumhansl, 1990). Finally, this finding fits well with the notion (paralleling other perceptual and cognitive domains) that the stable tones function as cognitive reference points in terms of which other tones are encoded and remembered.

Similar temporal-order effects also have been found in memory for tones (Krumhansl, 1979; Dowling & Bartlett, 1981; Bartlett & Dowling, 1988), relatedness judgments of chords (Krumhansl, Bharucha, & Kessler, 1982, Krumhansl, Bharucha, & Castellano, 1982; Bharucha & Krumhansl,
1983), and memory for chords (Krumhansl, Bharucha, & Castellano, 1982; Bharucha & Krumhansl, 1983). In all cases, the results can be accounted for in terms of differential stability of the elements in the underlying tonal and harmonic hierarchies, and the asymmetries mirror compositional practice (Krumhansl, 1990). It should be emphasized that, even though the contexts used in some of these studies may have been “static” scales and triads (used simply to instantiate a musical key), the cognitive representations revealed by the experiments carry important time-variant information.

Another principle concerning temporal-order relations is the melodic anchoring principle proposed by Bharucha (1984a). The principle describes the conditions under which an unstable tone “resolves” to a stable tone that is close in pitch by stepwise (immediate or delayed) motion along the chromatic or diatonic scales. In an extensive series of experiments, Bharucha demonstrated that this principle governs both the initial activation of tonal schemas (including finding a tonal center), and memory accuracy for tones in melodies once a tonal schema has been activated. Among other results, these studies showed reordering of tones produced clear shifts in tonal interpretations as predicted by the melodic anchoring principle.

Given the studies reviewed here, it seems inaccurate to describe the situation as does Butler (1989, p. 234, footnote 11): “The suggestion that compositional conventions regarding the time-ordering of ‘tendency’ tones might be central to psychological theories of harmony and tonality was not widely accepted within the community of music perception researchers in 1980 [when Butler presented his 1980 paper], and does not enjoy much wider acceptance now.” Rather, it would seem more apt to characterize the situation as one in which temporal-order effects have been rather extensively investigated and documented experimentally. However, important problems remain to be addressed concerning how pitch information is presented temporally to yield the sense of melodic and harmonic tendencies fundamental to the experience of music. Important issues also reside in the relationship between pitch ordering and metrical and rhythmic patterns. The experimental evidence Butler offers suggests that the “theory of intervalic rivalry,” with its narrow focus on rare intervals (the tritone, in particular), promises little progress toward these goals. 1

References


1. I am grateful for the helpful suggestions made on an earlier version of this paper by Lola Cuddy, James Cutting, Fred Lerdahl, and Johan Sundberg.


