Cognitive Foundations of Musical Pitch

CAROL L. KRUMHANSL
Department of Psychology
Cornell University

OXFORD PSYCHOLOGY SERIES NO. 17

New York Oxford
OXFORD UNIVERSITY PRESS
1990

4. A key-finding algorithm based on tonal hierarchies

The last chapter showed that the experimentally measured tonal hierarchies correlate strongly with the distribution of tones in tonal-harmonic music. This suggested that the tonal hierarchies might be acquired through experience with the musical style, particularly through internalizing the relative frequencies and durations with which tones are sounded. The question is turned around in this chapter. Here it is asked whether listeners could use the tonal hierarchies, once acquired, to determine the key of particular musical selections. The basic idea is that the tonal hierarchies function as a kind of template against which the tones of the musical selection are matched. This pattern-matching process is modeled by a computer algorithm—written in collaboration with Mark Schmuckler—which is applied to a variety of musical segments. To the extent that the key-finding algorithm produces correct results, it strengthens the case that pattern-matching to tonal hierarchies may be one mechanism through which listeners arrive at a sense of key.

A second motivation for developing and testing the algorithm was that determining key is prerequisite to successful automation of music analysis. The artificial-intelligence literature describes a number of attempts along these lines. These include harmonic analysis (Winograd, 1968; Winold & Bein, 1985) and Schenkerian analysis (Meehan, 1980; Smoliar, 1980) of tonal music, and set-class analysis of atonal music (Forte, 1970). For the automatic analysis of tonal music, the key needs to be determined in order for the structural roles of melodic and harmonic events to be coded meaningfully. For example, in connection with harmonic analysis, Winograd (1968) noted the inherent ambiguity of chords and the necessity to ascribe meaning to them in terms of their functions within the system of interrelated tonalities. Various approaches to the problem of determining the key of a piece are described in the artificial-intelligence literature.

Two algorithms, those of Longuet-Higgins and Steedman (1971) and Holtzman (1977), provide direct comparisons with the present algorithm, so they will be described in more detail later. Briefly, Longuet-Higgins and Steedman's algorithm (1971) successively eliminates keys on the basis of whether or not the tones in the musical sample are contained in the scale of each of the major and minor keys. Holtzman's algorithm (1977) examines the music for various key-defining features, such as the tonic triad (the major or minor chord built on the tonic), the tonic–fifth relationship (the interval between the tonic and the fifth scale tone, which is called the dominant), and the tonic–third relationship (the interval between the tonic and third scale tone, which is called the mediant and can be used to establish whether the key is major or minor). Both algorithms were quite successful when applied to the fugue subjects from Bach's *Well-Tempered Clavier.*

Sockut (described in Holtzman, 1977) devised a procedure for key-finding giving positive weights if the tonic, dominant, leading tone (seventh scale degree), and characteristic note sequences are contained in the music, and a negative weight if the tritone is present. Chafe, Mont-Reynaud, and Rush (1982) assigned key after previously finding points of rhythmic and melodic accent. Tones occurring at these points are searched for important pitch relations, and a list of possible tonics is formed. This list of tonics is weighted by frequency of occurrence and their own key implications, and from these statistics the best choice of key is found. Winold and Bein (1985), in their analysis of Bach chorales, used a somewhat similar approach by attaching weights to the tonics of all cadences, giving a double weight to the final cadence, and finding the most frequently occurring tonics. In their analysis, the cadences were determined externally to the algorithm by locating fermatas, which signal cadences in this style.

The algorithm described in this case differs from previous artificial-intelligence approaches to the problem of key-finding inasmuch as it is based directly on prior psychological results. An additional difference is that earlier approaches have taken as their objective the assignment of a single key, usually for the purpose of determining the appropriate key signature. As such, they treat musical key as a discrete, single-valued quality. In contrast, the present algorithm produces a vector of quantitative values. The key with the highest value might be taken as the main key; the extent to which it has been established is considered to be a matter of degree. In addition, the key assignment occurs within the context of other closely related, but subordinate keys. Thus, the algorithm produces a result that is consistent with the idea that at any point in time a listener may entertain multiple key hypotheses, which may or may not be made unambiguous by other features of the music. These hypotheses could serve to establish an appropriate tonal framework for encoding subsequent events and appreciating contrasts with other, more or less closely related keys.

The key-finding algorithm

As shown schematically in Figure 4.1, the input, **I**, to the algorithm is a 12-dimensional vector $\mathbf{I} = (d_1, d_2, \ldots, d_{12})$, specifying the total durations of the 12 chromatic scale tones in the musical selection to which the key is to be assigned. Note that the input vector might be based on a selection of any length, from a few tones to an entire composition. By convention

KEY-FINDING ALGORITHM

I = Tone durations in musical selection

$(d_1 \ldots d_{12})$

K_i = Probe tone profile for key i

$(P_{1i} \ldots P_{12i})$

R = Match between durations and profiles

$(r_1 \ldots r_{24})$

$$I \times \begin{matrix} K_1 \\ \vdots \\ K_{24} \end{matrix} \dashrightarrow R$$

Fig. 4.1. The key-finding algorithm takes as input the distribution of tone durations in the musical segment, **I**. It is correlated with the probe tone rating profile (Krumhansl & Kessler, 1982) for each of the 24 major and minor keys, $\mathbf{K}_i$. The output is a vector of numerical values indicating the strength of each key, **R**.

d_1 will specify the total number of beats that the tone C is sounded, d_2 will specify the total number of beats that the tone C♯ (or its enharmonic equivalent D♭) is sounded, and so on.

In the present applications, the durations in the input vector are specified in terms of numbers of beats but other units of time would do equally well. The input to the algorithm does not code the octave in which the tones are sounded or their order; how order information might be incorporated in an algorithm of this sort is discussed later. Nor does the algorithm distinguish between enharmonic tones: that is, tones played the same on fixed-pitch instruments but with different spellings in musical notation, such as C♯ and D♭. This means that the input to the algorithm might potentially come from a musical score, from a fixed-pitch instrument such as a keyboard, or from an analysis of acoustic information. However, once a key (or key region) has been determined, the correct spellings of the tones will be able to be determined in most cases. For example, if the predominant key contains a number of sharps, then the C♯ spelling would be preferred over D♭. If, on the other hand, the predominant key contains a number of flats, then the D♭ spelling would be preferred. Ambiguities would arise primarily in the region of F♯ or G♭ major, which have six sharps and six flats, respectively.

The algorithm correlates the input vector, **I**, with 24 stored 12-dimensional vectors, $\mathbf{K}_1, \mathbf{K}_2, \ldots, \mathbf{K}_{24}$. The 24 vectors represent the tonal hierarchies of the 24 major and minor keys. Each vector contains 12 values, which are the ratings from the Krumhansl and Kessler (1982) study, described in Chapter 2, of the degree to which each of the 12 chromatic scale tones fit with the particular key. The vector for C major is (6.35, 2.23, 3.48, 2.33, 4.38, 4.09, 2.52, 5.19, 2.39, 3.66, 2.29, 2.88). By convention, the first number corresponds to the tone C, the second to C♯ (D♭), and so on. The vector for D♭ major is found by shifting these values one place to the right and filling in the first value from the end of the C major vector, resulting in a D♭ major vector of (2.88, 6.35, 2.23, 3.48, 2.33, 4.38, 4.09, 2.52, 5.19, 2.39, 3.66, 2.29). The vectors for all other major keys are obtained by a similar procedure that is justified by the perceptual equivalence of major keys under transposition (Krumhansl & Kessler, 1982). The vector for C minor, again taken as the ratings from the probe tone experiment, is (6.33, 2.68, 3.52, 5.38, 2.60, 3.53, 2.54, 4.75, 3.98, 2.69, 3.34, 3.17). The vectors for all other minor keys, as for the major keys, are found by shifting the values the appropriate number of places to the right and filling in the resulting empty spaces by the values at the end of the C minor vector, which is justified by the perceptual equivalence of minor keys under transposition.

The input vector, **I**, is correlated with each of the vectors, $\mathbf{K}_1$, $\mathbf{K}_2, \ldots, \mathbf{K}_{24}$, producing an output vector of correlations, $\mathbf{R} = (r_1, r_2, \ldots, r_{24})$, in which r_i, is the correlation between the vectors, **I** and $\mathbf{K}_i$, computed as described in Chapter 2. In this instance, it is a measure of the degree to which the durations of the tones in the input segment match the tonal hierarchy of each key. The correlation will be high to the extent that tones with the longest relative durations match tones dominating in the tonal hierarchy. Each correlation can be evaluated for statistical significance. Note that the number of degrees of freedom for the correlation is constant; it does not depend on the size of the musical selection on which the input vector is based. Nor is the correlation affected by the mean or standard deviation of the values of the input vector (which would vary with sample size). Hence, the algorithm can be applied to musical selections varying over a wide range of lengths, although sections containing only a few tones would probably produce unstable results.

Further analyses of the output vector, **R**, can be performed depending on the application. If the objective is to assign a key signature to a composition as a whole, then the input vector might be based on an initial segment of the piece, on a final segment, or even on the entire piece; in this case, the highest value in the output vector would determine the best choice of key signature based on the sample segment. If the objective is to detect possible changes of key within the composition, then the algorithm might use as input the durations of tones within significant subsections of the composition, or within initial or final segments of these subsections. In both of these applications, the magnitude of the highest correlation (or its level of significance) would be expected to be high and the key assignment relatively unambiguous. If the objective is to trace shifting tonal orientations or temporary modulations, then the algorithm might move through an entire piece using as input the tone durations for segments of only a measure or two in length. For some segments, it may be that no single correlation is very high and a number of keys have approximately equal values. This ambiguity might be represented as a point between keys in a spatial map; one method for doing this is employed here.

Three applications test this simple algorithm for determining key strengths. In the first, the algorithm is applied to the initial four-note segments of Bach's 48 preludes from the *Well-Tempered Clavier,* and the results are compared with those of a study by A. J. Cohen (1977) in which she asked listeners to identify the key of a subset of the preludes after hearing short initial and final segments. To investigate the robustness of the algorithm across musical styles, the same analysis is performed on the 24 preludes of Shostakovich and Chopin. In the second application, the algorithm is applied to the fugue subjects of Bach's 48 fugues in the *Well-Tempered Clavier.* In this application, the input sample is increased in length until it includes the entire subject. This application determines the first point at which the correct key is assigned, and the results are compared with those of Longuet-Higgins and Steedman (1971) and Holtzman (1977). The same kind of analysis is done of the Shostakovich fugues. In the final application, one of Bach's preludes containing an interesting pattern of shifting tonal centers is analyzed by the algorithm on a measure-by-measure basis. The results are compared to analyses made by two music theory experts who were asked to assign a most likely key interpretation to each measure and indicate other keys of lesser strengths.

Application I: initial segments of preludes of J. S. Bach, Shostakovich, and Chopin

The key-finding algorithm was first applied to the 48 preludes of Bach's *Well-Tempered Clavier.* This choice was made for two reasons. First, these preludes (and their accompanying fugues) move through the entire set of 24 major and minor keys and were written to demonstrate the utility of equal-tempered tuning, which allows any tonality to be played on fixed-pitch instruments like the piano. Each prelude begins quite unambiguously in the key of the key signature; therefore failures of the algorithm would suggest that it should be abandoned immediately. Second, A. J. Cohen (1977) described a study in which listeners, who were all university music majors, heard short segments of 12 of the preludes. They were then asked to sing the scale of the key in which they thought the piece was written. In one condition of her experiment, listeners heard just the first four sounded events and were quite accurate in their responses, choosing the correct key about 75 percent of the time. Applying the algorithm to the preludes allows us to compare its performance to these psychological data.

The input vectors were based on the first four notes of the preludes. Figure 4.2 gives as an illustration the initial segment of the C Minor Prelude, Book I, and indicates the four notes on which the input vector **I**, was based. The tone C is sounded for two sixteenth notes (giving a total of one half beat); the E♭ and G are each sounded for one sixteenth note (one quarter beat each); all other tones have zero duration in this sample. This results in an **I** vector of (.5, 0, 0, .25, 0, 0, 0, .25, 0, 0, 0, 0). For some of the preludes, the four-tone rule had to be modified slightly because two or more tones sounded simultaneously at the time when a fourth tone was sounded; all the simultaneously sounded tones were then included in the sample. The input vectors used the full notated temporal durations of the tones in the sample until the time at which either the fourth tone ended, or the fifth tone began.

The input vectors for each of the 48 preludes were correlated with the 24 tonal hierarchy vectors, $\mathbf{K}_i$. Let d denote the designated (intended) key, and r_d the correlation of the input vector with the tonal hierarchy of the designated key. The value r_d is considered to be the strength of the in-

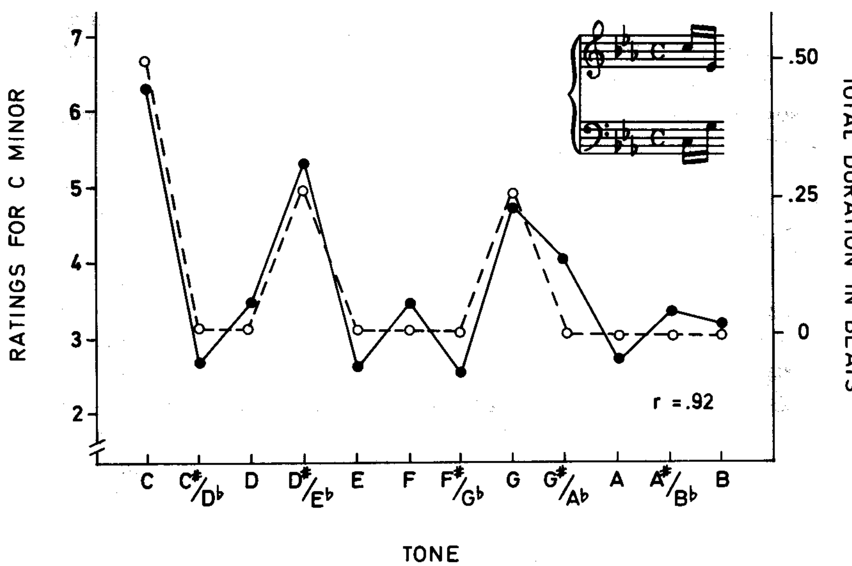

Fig. 4.2. The durations of the first four tones in J. S. Bach's C Minor Prelude, Book I, are plotted together with Krumhansl and Kessler's probe tone ratings for C minor (1982). The durations correlate significantly, r = .92, with the probe tone ratings for the intended key. The algorithm finds C minor to be the strongest key.

tended key as established by the input segment. A number of questions can be asked about the r_d value. First, we can ask if the value is statistically significant; if so, then the correspondence between the distribution of durations and the probe tone ratings for the intended key is unlikely to have occurred by chance. Second, we can ask if the r_d value is the largest in the output vector; if so, then we can say that the algorithm has found the correct key. Third, if the r_d value is not the largest, then we can ask how many other keys have higher r values and how far these other keys are from the correct key. Figure 4.2 plots, for the C Minor Prelude, the duration values of the first four tones superimposed on the tonal hierarchy of the intended key; the r_d value is .92, which is statistically significant and the highest value in the output vector, **R**.

Table 4.1 shows, for each of the 48 preludes, the r_d values based on the first four tones; each of these values was statistically significant, as indicated by the asterisks in the table. They averaged .83 and .79 for Books I and II, respectively, and .77 and .85 for major and minor preludes, respectively. The overall average was .81. In all but four cases, the r_d value was also the largest in the output vector. Table 4.1 indicates for the four exceptional cases the number of other keys with higher r values and their average distance from the intended key in Schoenberg's maps of key regions (1954/1969) shown in Figure 2.9. Distance is measured in terms of the number of keys along vertical and/or horizontal dimensions using a city-block metric. In the four cases in which some other key or keys had higher r values, these keys are closely related to the intended key with an average of 1.40 steps away. In two cases, the algorithm incorrectly assigns the parallel major of the designated key; in both, the third scale degree (which differentiates the modes) is missing from the input segment. Example 1 in Figure 4.3 illustrates this with the input segment of the C Minor Prelude (Book II), to which the algorithm assigns C major. In the remaining two cases, the algorithm assigns the relative minor of the key of the dominant. To illustrate this, Ex. 2 shows the input segment of the E♭ Major Prelude (Book II), to which the algorithm assigns G minor because of the relatively long G and B♭ with which the prelude begins. In general, however, the algorithm was quite accurate, finding the correct key in almost all cases and identifying the correct key region in the remainder.

Table 4.2 compares the results of our algorithm with those of the listeners in A. J. Cohen's experiment (1977) for the 12 Bach preludes used in her study. The stimuli used the first four events, or attack points, of the preludes, which meant (because of simultaneously sounded tones) that more than four tones were included in some cases. The r_d values shown in the table were computed in two different ways: first, using the tones actually sounded in Cohen's experiment and, second, using the first four tones, as before. Comparison of the last two columns shows substantial agreement between the r_d values, however. In all cases, they were significant and, for this sample of preludes, they were always the highest

Table 4.1. Application of the key-finding algorithm to the initial segments of the Bach preludes

	Book I		Book II	
Prelude	r_d	Other Keys[a]	r_d	Other Keys[a]
1 C major	.81*		.69*	
2 C minor	.92*		.73*	1 (1)
3 C♯ major	.83*		.82*	
4 C♯ minor	.87*		.89*	
5 D major	.73*		.69*	
6 D minor	.81*		.68*	
7 E♭ major	.87*		.60*	2 (1.5)
8 D♯/E♭ minor	.92*		.77*	
9 E major	.83*		.81*	
10 E minor	.92*		.78*	
11 F major	.83*		.66*	
12 F minor	.85*		.90*	
13 F♯ major	.67*	1 (2)	.90*	
14 F♯ minor	.93*		.89*	
15 G major	.83*		.83*	
16 G minor	.85*		.75*	1 (1)
17 A♭ major	.87*		.83*	
18 G♯ minor	.83*		.92*	
19 A major	.73*		.79*	
20 A minor	.82*		.84*	
21 B♭ major	.88*		.66*	
22 B♭ minor	.91*		.77*	
23 B major	.68*		.73*	
24 B minor	.83*		.92*	

[a]The first number shows the number of keys with higher r values than r_d; the number in parentheses indicates their average distance from the intended key in Schoenberg's maps of key regions (1954/1969).

*Significant at $p < .05$.

in the output vectors. If cases in which the r_d value is the largest are counted as correct responses, then our algorithm was correct 100 percent of the time, whereas Cohen's listeners were correct on average only 75 percent of the time.

A number of further analyses were performed to determine whether the algorithm modeled the variations in performance across preludes exhibited by A. J. Cohen's listeners (1977). In Cohen's study, listeners performed better on major than minor preludes, which may indicate a perceptual bias toward the more stable major mode. In contrast, the algorithm shows r_d values for minor keys as higher on average than those

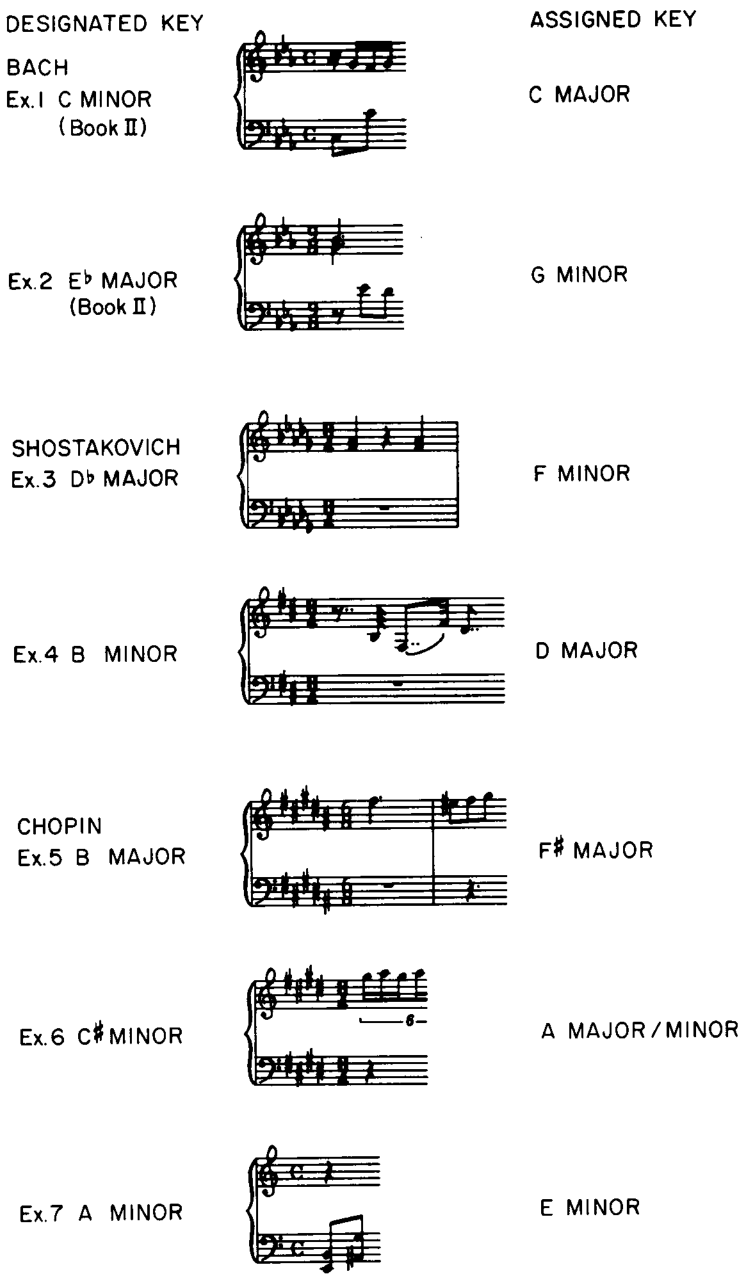

Fig. 4.3. Examples in which the present algorithm fails to find the correct key for four-tone initial segments of preludes by Bach, Shostakovich, and Chopin. The designated key and the key assigned by the algorithm are indicated in each case.

Table 4.2. Comparison of the key-finding algorithm with Cohen's listeners (1977) for preludes in Book I

		Computation of Strength of Intended Key[a]	
Prelude	Percentage Correct	$r_d 1$	$r_d 2$
1 C major	.89	.81*	.81*
3 C♯ major	.67	.79*	.83*
5 D major	.72	.73*	.73*
7 E♭ major	.89	.87*	.87*
9 E major	.94	.81*	.83*
11 F major	.89	.78*	.83*
Average	.83	.80*	.82*
Minor Preludes			
2 C minor	.56	.87*	.92*
4 C♯ minor	.83	.84*	.87*
6 D minor	.72	.83*	.81*
8 E♭ minor	.67	.92*	.92*
10 E minor	.56	.93*	.92*
12 F minor	.61	.88*	.85*
Average	.66	.88*	.88*

[a] $r_d 1$ is based on durations of tones used in A. J. Cohen's stimuli (1977). $r_d 2$ is based on durations of first four tones.

*Significant at $p < .05$.

for major keys. For the major and minor keys considered separately, no correspondence was found between the listeners' performance and the values of r_d. Nor was a correspondence found between the listeners' performance and the difference between this value and the next highest r value in the output vector, which might be taken as a measure of tonal ambiguity. Thus, the algorithm does not mirror in a detailed way the variations in level of performance across preludes, although it is unclear whether these differences are statistically reliable in Cohen's study. In addition, the lack of correspondence may be due to the algorithm performing at a ceiling level, whereas performance was lower in Cohen's study possibly because of limits in memory or response production.

In tribute to J. S. Bach, Shostakovich wrote a set of 24 preludes and fugues that were completed in 1951. Despite their relative recency, these works are highly tonal and, like the *Well-Tempered Clavier,* they move completely through the set of 24 major and minor keys. To determine the algorithm's success at finding the intended keys of these preludes, the same kind of analysis was carried out using as input samples the first four tones; the results are shown on the left of Table 4.3, where the preludes are reordered to correspond to the Bach preludes.

The analysis of the Shostakovich preludes found that in all but three cases the r_d value, measuring the strength of the intended key, was statistically significant and, in all but seven cases, it was the highest in the output vector, **R**. The r_d values averaged .71 and .75 for major and minor preludes, respectively, with an overall average of .73. The other keys with higher r values than r_d were, however, relatively close to the intended key, with an average distance of 1.42 steps. In four of the seven cases, the algorithm confuses parallel major and minor keys; in each case, the third scale degree is missing from the input segment. In two other cases, the input segment consists of only one or two different tones and the algo-

Table 4.3. Application of the key-finding algorithm to the initial segments of the Shostakovich and Chopin preludes

	Shostakovich		Chopin	
Prelude	r_d	Other Keys	r_d	Other Keys
1 C major	.80*		.81*	
20 C minor	.71*	1 (1)	.88*	
15 D♭ major	.46	2 (1.5)	.82*	
10 C♯ minor	.83*		.25	6 (1.67)
5 D major	.88*		.27	3 (1)
24 D minor	.87*		.76*	
19 E♭ major	.72*		.59*	2 (1.5)
14 E♭ minor	.27	4 (1.75)	.71*	1 (1)
9 E major	.67*		.88*	
4 E minor	.74*	1 (1)	.55	2 (1.5)
23 F major	.80*		.76*	
18 F minor	.89*		.00	11 (2.45)
13 F♯ major	.64*		.88*	
8 F♯ minor	.83*		.38	4 (1.25)
3 G major	.57	1 (1)	.79*	
22 G minor	.85*		.21	7 (2.71)
17 A♭ major	.84*		.76*	
12 G♯ minor	.84*		.85*	
7 A major	.67*	1 (1)	.49	3 (1.67)
2 A minor	.74*		−.08	11 (2.55)
21 B♭ major	.79*		.53	2 (1.5)
16 B♭ minor	.89*		.18	6 (1.83)
11 B major	.72*		.38	3 (1.67)
6 B minor	.59*	2 (1.5)	.92*	

*Significant at $p < .05$.

rithm's assignments are reasonable given the inadequate information. Example 3 in Figure 4.3 shows the input segment for the D♭ Major Prelude to which the algorithm understandably assigns F minor (in which the two sounded tones are the tonic and third scale degree). The final case, the B Minor Prelude shown in Ex. 4, is more problematic. The initial segment contains the tonic triad of the key, but because of the long durations of the third and fifth scale degrees, the algorithm incorrectly assigns D major. Overall, however, the algorithm was quite accurate, determining the correct key in most cases and the correct key region in all cases.

The final analysis in this section is of the 24 Chopin preludes, also composed in homage to Bach. However, these preludes differ from those of Bach and Shostakovich in their more pronounced tonal ambiguity, the use of expanded diatonic and chromatic vocabulary, and the roles of dissonances and chromaticism. They are of interest, therefore, for exploring the range of musical styles to which the algorithm might be applied. The results of the analysis using input vectors based on the first four tones are shown on the right of Table 4.3. In 13 cases, the r_d value was statistically significant and in 11 cases the intended key had the highest value in the output vector. The average r_d values were .66 and .47 for major and minor preludes, respectively, with an overall average of .57

Thirteen Chopin preludes had one or more keys with r values higher than r_d; their average distance from the intended key was 2.02 steps. These values are considerably larger than in previous applications. The 13 cases can be classified into three groups. The first group contains eight preludes for which four or fewer keys had higher values than the intended key. These keys were an average of 1.40 steps away from the intended key, a distance comparable to that for the Bach and Shostakovich preludes. In one case the algorithm found the relative major of the designated key, in one case the key of the subdominant, and in six cases the key of the dominant. Example 5 in Figure 4.3 shows one of the latter cases, the Prelude in B Major. The input segment clearly emphasizes F♯ and as a consequence F♯ major was the assigned key. In all these cases, however, the algorithm is finding the correct key region.

The second group contains the three preludes in C♯ minor, G minor, and B♭ minor for which six or seven keys had higher values than the intended key, with an average distance of 2.11 steps. In none of these cases was an r value very large or significant (the highest value was .44). In all three cases, the tones in the input sample are contained in the scale of the intended key, but the tonic is absent and at most one of the tones of secondary importance (third and fifth scale degrees) appears. In the face of this, the algorithm determined that no key is very strongly indicated by the initial segments. Keys with higher r values than the intended key tended to be located in a diffuse region around the intended key. Example 6 shows the input segment for the Prelude in C♯ Minor to which the algorithm assigned weak A major and A minor keys (in which the tones are leading tone and tonic).

The two final cases, the Preludes in F Minor and A Minor, have strikingly low r_d values and 11 other keys with higher values the distances of which average 2.50 steps away from the intended key. Example 7 shows the input segment for the A Minor Prelude to which the algorithm assigned E minor. In this segment, the tonic of the designated key is raised by a semitone from A to A♯. This means that the input segment contains the tonic, third, and fifth scale degrees of E minor. The Prelude in F Minor is similar; it contains a chromatic alteration of the third scale degree, resulting in a better match to B♭ minor. Note that in both these cases, however, the key found to be strongest by the algorithm is an immediate neighbor of the correct key.

To summarize, these applications used as the input vector the first four tones of each prelude in the three sets. It was most successful in finding the keys of Bach's preludes. In the few exceptional cases, a key close to the intended key was selected instead. Comparison with A. J. Cohen's experiment (1977) showed that the algorithm was more accurate than her subjects; also, the algorithm did not mirror the specific pattern of errors across the subset of preludes used in the experiment. In the analysis of the Shostakovich preludes, the algorithm selected the intended key in somewhat fewer cases, but in these the correct key region was always determined. The algorithm had more difficulty with the Chopin preludes, but still performed far better than the chance level of 1 in 24. More important, in the large majority of cases it identified the correct key region, and errors could be traced directly to the input segments themselves. It should be emphasized that the segments consisted of just four tones, which in quite a few cases contained duplicated tones. Given this, it was encouraging that the results were as accurate as they were. This raises the question, however, of what would happen if the input segments were increased in length. The next application addresses this question.

Application II: fugue subjects of J. S. Bach and Shostakovich

Longuet-Higgins and Steedman (1971) proposed a computer algorithm for assigning key, which they applied to the fugue subjects from Bach's *Well-Tempered Clavier.* This algorithm matches the tones in the fugue subjects to box-shaped regions delimiting the scale tones of the key in a spatial array. This array, shown in Figure 4.4, is such that neighbors in the horizontal dimension are separated by an interval of a fifth and neighbors in the vertical dimension are separated by an interval of a major third (similar arrays can be found in Fokker, 1949; Helmholtz, 1885/1954). Each major and minor key corresponds to a set of diatonic scale tones forming a compact set in this array. The sets for C major and C minor are shown

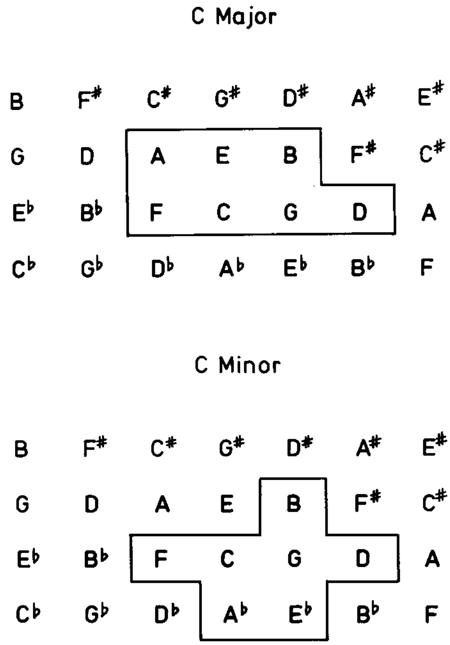

Fig. 4.4. The arrays of scale tones for C major (top) and C minor (bottom) from Longuet-Higgins and Steedman's key-finding algorithm (1971). Their algorithm, applied to the fugue subjects of Bach's *Well-Tempered Clavier,* works by successively eliminating keys the arrays of which do not contain the tones of the fugue subject. Adapted by permission of the author and publisher.

in the figure; the set for any other major or minor key can be found by shifting the box-shaped region appropriately.

Their algorithm works by eliminating musical keys as the fugue subject progresses. The first note eliminates those keys in the box-shaped region of which the first tone is not contained. Then, the second tone eliminates from the remaining keys those in which the second tone is not contained. This process is continued until one single key remains, and the step (number of tones required) at which this occurs is recorded. Applied to the fugue subjects, the correct key is found in 26 of the 48 cases.

The remaining cases are of two types. In the first, a conflict arises at some step in the process so that all keys have been eliminated, that is, no single key contains all the tones. When this state of affairs occurs, an additional rule, the tonic–dominant preference rule, is invoked. The algorithm goes back to the point just before the conflict occurred, and first preference is given to the key the tonic of which is the first tone of the fugue subject if this was one of the candidate keys. If not, the rule gives second preference to the key the dominant (fifth scale tone) of which is the first tone of the fugue subject. There were fives cases in which the tonic–dominant preference rule was invoked when a conflict of this sort arose, giving the correct key in all cases.

The second kind of case is one in which the end of the fugue subject is reached and two or more possible keys remain. That is, the set of pitches in the fugue subject is not unique to a single key. At this point, the tonic–dominant rule is also invoked; there were 17 cases in which the correct key was found in this way. Table 4.4 shows the step at which Longuet-Higgins and Steedman's algorithm (1971) found the correct key, that is,

Table 4.4. Application of the key-finding algorithm to the fugue subjects of the Bach fugues

	Book I			Book II		
Fugue	r_d	Step	L-H&S[a]	r_d	Step	L-H&S
1 C major	.51	2	16+	.79*	4	23+
2 C minor	.70*	5	5	.70*	5	9+
3 C♯ major	.82*	7	16	.67*	2	4+
4 C♯ minor	.74*	3	4	.59*	12	12
5 D major	.51	2	15+	.84*	10+	9+
6 D minor	.68*	3	8	.68*	3	15
7 E♭ major	.71*	6	11+	.83*	2	20+
8 D♯ minor	.69*	6	12+	.61*	15++[c]	9
9 E major	.63*	12+[b]	11	.64*	2	6+
10 E minor	.83*	2	7+	.73*	15	18
11 F major	.60*	10	6	.74*	4	17
12 F minor	.62*	15	4+	.77*	5	7
13 F♯ major	.81*	2	8	.55	5	12+
14 F♯ minor	.71*	18	5+	.89*	3	18+
15 G major	.64*	2	15	.76*	4	16
16 G minor	.66*	3	4	.33	18+	18+
17 A♭ major	.81*	2	7+	.69*	12	22+
18 G♯ minor	.75*	5	5	.68*	3	25+
19 A major	.64*	4	7	.51	2	20+
20 A minor	.72*	5	5	.64*	9	5
21 B♭ major	.68*	4	14	.51	2	9
22 B♭ minor	.72*	3	6+	.62*	3	5
23 B major	.71*	11	11	.67*	2	12+
24 B minor	.89*	3	7	.89*	3	6

[a]Longuet-Higgins and Steedman's algorithm.

[b]The symbol + indicates that the tonic–dominant preference rule was invoked and selected the intended key.

[c]The symbol ++ indicates that the tonic–dominant preference rule was invoked and did not select the intended key.

the number of tones needed to find the key. The table also indicates those cases requiring the tonic–dominant preference rule.

To compare their algorithm with our own, we again used the fugue subjects. A duration profile was found for samples varying in length from one tone to the entire subject. This means that, for each of the fugues, there were n input vectors, $\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n$, where n is the number of tones in the fugue subject. Each of these was correlated with the tonal hierarchies of the 24 major and minor keys producing n output vectors, $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_n$. The first step (the number of tones in the input sample) for which r_d is the highest in the output vector was recorded. We say this is the step at which the algorithm finds the correct key.

Figure 4.5 illustrates the application of the algorithm to two fugues, in C major and F major, from Book I. For the C Major Fugue (Fugue I), the first input sample contains the tone C, the second C and D, the third C, D, and E, and so on. The first point at which the algorithm finds that C major has the highest r value in the output vector is after the first two tones. For the Fugue in F Major (Fugue XI), the correct key is not determined until after the tenth tone. Until that time, C major has consistently higher values despite the early sounding of the B♭ (which is the only tone distinguishing between F and C major). This can be understood because the tone C is emphasized more than the tone F at the beginning of the subject. Only after the second repetition of the B♭ is the algorithm able to find the correct key.

Table 4.4 shows, for comparison with the Longuet-Higgins and Steed-

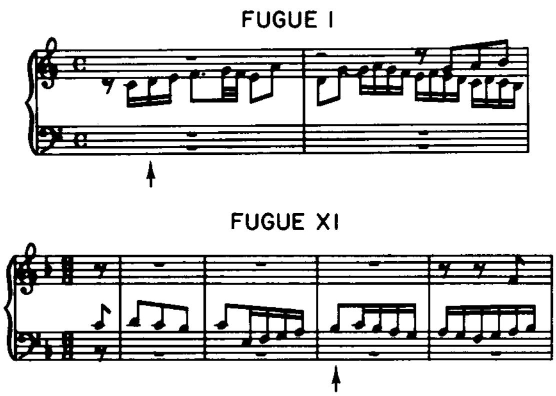

Fig. 4.5. The fugue subjects from the C Major Fugue and the F Major Fugue from Bach's *Well-Tempered Clavier* (Book I). In the first case, the present algorithm (which matches tone durations to tonal hierarchies) determines the correct key after the first two tones; that is, C major is the strongest key in the output vector when two tones are included in the input vector. In the second case, 10 tones are required to find the correct key of F major.

man (1971) algorithm, the step at which our algorithm finds the designated key for each of the 48 fugues, together with the corresponding r_d values. In all but six cases, the r_d value, measuring the strength of the intended key, was statistically significant at the point when it was first maximum, averaging .69 across Books I and II. In four cases, the algorithm did not find the correct key by the end of the fugue subject. In three of these, the tonic–dominant preference rule would eliminate all keys having higher r values at the end of the fugue subject. For the remaining case, the Fugue in D♯ Minor, the parallel major of the intended key had a higher correlation at the end of the fugue subject. Because this key has the same tonic and dominant as the designated key, the rule cannot be used to find the correct key.

The number of tones required to determine the intended key is of primary interest for comparing the two algorithms. Figure 4.6 shows, for the two algorithms, the distribution of steps at which the correct key was found for those fugues to which the tonic–dominant preference rule was not applied. It is clear that the present algorithm found the correct key considerably faster, requiring only two or three tones in almost half the cases. The average number of tones needed was 5.11 for the present algorithm and 9.42 for that of Longuet-Higgins and Steedman (1971).

We believe the reason for the more efficient performance of the present algorithm is that it makes finer gradations among the tones by assigning them numerical values in a tonal hierarchy. Thus, the pattern-matching process is one in which the tones are weighted by their tonal significance. In contrast, the algorithm proposed by Longuet-Higgins and Steedman (1971) is based on the distinction between scale and nonscale tones without considering the relative importance of the scale tones in the key. In fairness, though, it should be noted that their approach has the advantage that the stopping rule is internal to the algorithm; the algorithm proceeds until all but one key has been eliminated or the end of the fugue subject has been reached. In our approach, we used prior knowledge of the intended key and examined the output of the algorithm for the point at which this was the key of maximal strength. At subsequent points in time, other keys will likely have higher values. This, however, reflects the common practice of shifting to related keys in tonal music, a point that is directly addressed in the application in the next section.

Holtzman (1977) also applied his algorithm to the Bach fugues, using Book I only. The entire fugue subject was employed in each case. First, the algorithm examines the subject for the presence of a tonic triad chord. If one is not present, then it searches for the tonic–fifth relationship and the tonic–third relationship (which can be used to establish whether the key is major or minor) at significant points in the music, such as the first two tones of the subject and the first and last tones. Failing this test, the music is examined at other points for these relationships. If the tests still yield ambiguous results, then the algorithm proceeds to a recursive pro-

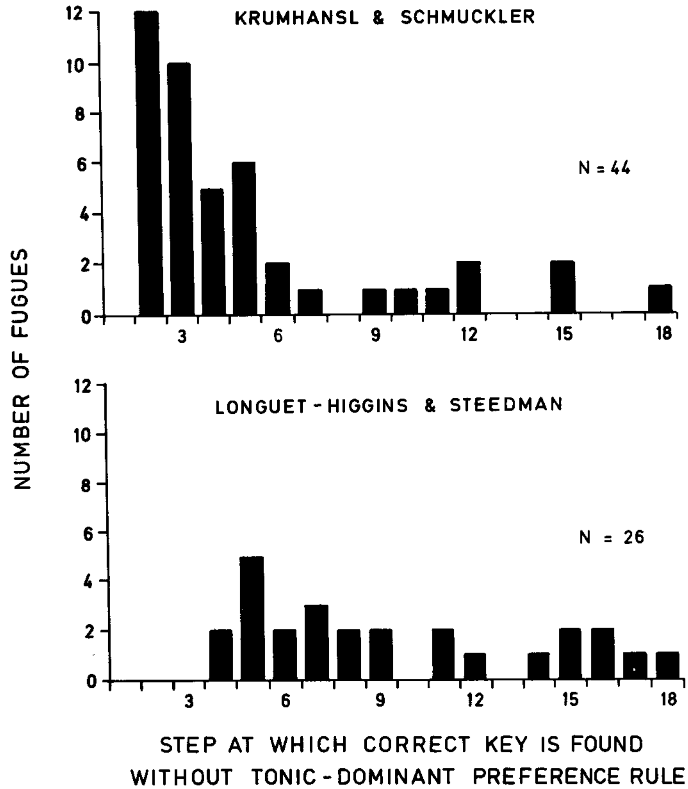

Fig. 4.6. The step (number of tones) at which the present algorithm (top) and the algorithm of Longuet-Higgins and Steedman (1971) find the correct key for the Bach fugues. These are all cases in which the tonic–dominant preference rule is not invoked. The average number of tones needed is 5.11 for the present algorithm and 9.42 for the earlier algorithm.

cedure. In this, all possible keys consistent with the tones are determined, weighted according to whether the tonic and dominant tones are present. If no conclusive decision can be reached with the whole input string, then the last tone is removed and the process repeated until a key is determined or fewer than six tones remain, in which case the algorithm concludes that the music is chromatic (i.e., not clearly in any key). Holtzman's algorithm differs from that of Longuet-Higgins and Steedman (1971) in that, whereas both look for scale tones, the former allows a small number of nondiatonic tones.

In Holtzman's application (1977) of this algorithm to the 24 fugues in Book I, it determined the correct key in all but one case. This performance rate is identical to that of the present algorithm (without the use of the tonic–dominant preference rule) for these fugues. In one case, the Fugue in F♯ Minor, Holtzman's algorithm was unable to find the key, and this is the fugue requiring the longest input segment for the present algorithm. The similarity between the results is probably due to both algorithms emphasizing the importance of the tonic, third, and fifth scale tones for defining the key. The present algorithm, however, is simpler computationally and determines the correct key with input segments considerably shorter than the entire fugue subjects used by Holtzman.

The final analysis presented here is of the subjects of the Shostakovich fugues, which, like the preludes, are written in a highly tonal style. Table 4.5 shows the r_d values and the step at which the intended key first had a maximal value. All but two of the correlations at these steps were statis-

Table 4.5. Application of the key-finding algorithm to the fugue subjects of the Shostakovich fugues

Fugue	r_d	Step[a]
1 C major	.75*	3
20 C minor	.59*	2
15 D♭ major	.21	21+
10 C♯ minor	.72*	13++
5 D major	.71*	7
24 D minor	.75*	15
19 E♭ major	.74*	2
14 E♭ minor	.65*	22
9 E major	.68*	2
4 E minor	.84*	2
23 F major	.73*	3
18 F minor	.84*	2
13 F♯ major	.67*	2
8 F♯ minor	.66*	9
3 G major	.51	2
22 G minor	.68*	3
17 A♭ major	.81*	3
12 G♯ minor	.64*	2
7 A major	.80*	2
2 A minor	.67*	3
21 B♭ major	.83*	2
16 B♭ minor	.68*	2
11 B major	.67*	2
6 B minor	.78*	3

*Significant at $p < .05$.

[a]The symbol + indicates that the tonic–dominant preference rule was invoked and selected the intended key; the symbol ++ indicates that the tonic–dominant preference rule was invoked and did not select the intended key.

tically significant, with an overall average of .69. Furthermore, in all but two cases, the algorithm found the correct key without invoking the tonic–dominant preference rule. Excluding these last two cases, the average number of tones needed was 4.32, and in 18 of the 24 cases the algorithm found the correct key with three tones or less. For one of the other two cases, the tonic–dominant preference rule would eliminate keys with higher correlations than the intended key. In the remaining case, the parallel major of the intended key, which has the same tonic and dominant, had a higher correlation at the end of the fugue subject; therefore the rule could not be used to find the correct key. In general, however, the algorithm proved efficient in finding the keys of the Shostakovich fugues, just as it had for the Bach fugues.

Application III: J. S. Bach's C Minor Prelude, Book II

In the third and final application of the key-finding algorithm, we examined the entire Prelude in C Minor from Book II of Bach's *Well-Tempered Clavier.* We selected this prelude because it contains an interesting pattern of shifting tonal centers, and we were interested in knowing whether the algorithm would be sensitive to this aspect of the piece. To have a basis for comparison, two music theorists were independently asked to provide an analysis of the prelude on a measure-by-measure basis in a form comparable to the output vector of the algorithm. We chose one-measure units because initial explorations of the algorithm showed that they provided the clearest pattern of shifting tonal centers while maintaining a fairly high degree of stability. The experts were asked to indicate the primary key for each measure and also any keys of lesser strengths. Thus, this application examines the extent to which the algorithm models experts' judgments about the keys suggested at a local level by this composition.

More specifically, the experts were asked to rate on a 10-point scale the strengths of any keys suggested by each of the 28 measures of the prelude. If a key was judged to be at maximal strength, then it was to be given a rating of 10; if, however, a key was only very weakly suggested, then it was to be given a rating of 1. It was understood that the majority of "no-mention" keys were implicitly given a rating of zero. Both of the experts had previous familiarity with the piece, had at their disposal the score and a piano so that they could play the prelude, and had a variety of analytical tools that they were free to employ. Each provided a detailed account of the structure of the piece in addition to the numerical ratings.

The first expert, Daniel Werts, was asked to consult on the project because his theory of scale reference (Werts, 1983) provides a framework that is especially congenial to the present approach. In the theory, a primary scale reference, or key, is normally established through some combination of the following means: emphasis of the tonic triad pitches in the outer voices, durational emphasis, presence of a cadence, and presence of a motif. A primary scale reference may be strong and unambiguous or it may be weak, such as during transitions, development sections, or other tonally unstable sections.

In addition, the theory allows for the possibility of secondary and tertiary scale references. Secondary scale references are described as projected by harmonic means, in which a chord borrowed from one scale can be injected into a context controlled by another, for example, an applied dominant. Tertiary scale references occur through nonharmonic means to three main ends: to foreshadow and echo primary and secondary references, to create melodic interest, or to avoid harmonically unacceptable phenomena. An example of a tertiary scale reference is the use of chromatic neighbors.

Basic to Werts's theory (1983) is the notion that keys are established to varying degrees that, in principle, would be possible to quantify and compare with the results of the key-finding algorithm described here. A second point of contact between the two approaches is the expression of key relations in the form of a toroidal configuration. Figure 2.10 shows his representation, which has the same basic structure as the multidimensional scaling of the correlations between the tonal hierarchies.

Werts (personal communication, June 1985) provided us with a graph, using a scale from 1 to 10, of the strength to which various keys are heard at each point throughout the prelude, and noted the advantages of construing modulatory strengths as points along a continuum:

> The flexibility of this method should be immediately apparent: one can depict not only differences in key-strengths, but also differences in the *rates* at which keys ebb and flow; moreover, the simultaneous presentation of several keys, which poses severe problems in most theories of modulation, submits readily to graphing.

The numerical values shown in the graph were justified by detailed observations concerning the use of chromatic tones, harmonic progressions, and adjacency in Werts's key scheme (1983). The graphed values were then integrated over units of one measure in length, the size of the unit of analysis for this application. These values are shown in Table 4.6 in rank order of their relative strengths. Because some keys were judged as exerting an influence for a period of less than one full measure, some of the values shown in the table are less than one. The number of keys receiving a nonzero value in a measure ranged from one to four, with an average of 1.93, indicating that more than one key was judged to be heard simultaneously in many measures of the piece.

The second expert, Gregory Sandell, was asked to assess key strengths independently on a measure-by-measure basis for the same Bach prelude. After graduate training in music theory at the Eastman School of Music, he came to Cornell University where he was involved in a number of

Table 4.6. Judgments of key strengths (Expert 1)[a] for Bach's C Minor Prelude (Book II)

Measure	Key 1	Key 2	Key 3	Key 4
1	c 7.58			
2	c 10.00			
3	c 10.00			
4	c 9.42	B♭ 3.81		
5	c 9.76	B♭ .36		
6	c 8.07	E♭ 4.67	A♭ .59	
7	E♭ 7.78	c 4.33	A♭ .89	
8	E♭ 9.25			
9	E♭ 10.00			
10	E♭ 10.00			
11	E♭ 10.00			
12	E♭ 10.00			
13	E♭ 10.00	B♭ 1.63		
14	E♭ 8.39	c 5.49		
15	c 8.11	E♭ 4.67	A♭ 1.17	
16	f 5.00	A♭ 2.89	c 1.78	
17	f 8.50			
18	f 9.50	b♭ .15		
19	f 10.00			
20	f 8.44	A♭ 5.72	E♭ .33	
21	f 8.77	b♭ 2.25	A♭ 1.29	
22	f 9.96			
23	f 9.11	E♭ 3.50		
24	E♭ 7.89	c 3.00	f .67	
25	c 8.04	E♭ 3.92		
26	c 9.65	E♭ 1.67		
27	c 9.46	f 1.61	E♭ 1.50	G .43
28	c 10.00	G .46		

[a]Lower-case letters indicate minor keys.

ongoing music perception projects. He was, however, unfamiliar with the nature of the key-finding algorithm. In addition to the numerical ratings requested, he provided a set of criteria for determining key strengths, and ranked them in terms of their importance (G. Sandell, personal communication, June 1985).

The first and most important criterion examines the music for the presence of a number of common-practice harmonic progressions: three-chord (or longer) expansions of the tonic, two-chord (or longer) expansions of the dominant or subdominant, cadential formulas ending on the tonic triad, metrically and durationally prominent half-cadences, and sequential contrapuntal or harmonic patterns. Such stylistically common progressions are considered to create expectations for specific continuations, thereby establishing specific keys. The second most important criterion, analogous to the key-finding algorithm of Longuet-Higgins and Steedman (1971), is that a key is a candidate if its set of scale tones is in operation over three successive harmonies. The third criterion concerns residual key effects, and asserts that the home key has an influence throughout the entire piece and that any other key established by a cadence has a continuing presence until approximately one measure beyond a subsequent cadence in a new key. The final criterion says that the simple presence of a major or minor triad makes the key of which it is the tonic a possible candidate, albeit a weak one on that basis alone. Its candidacy is more strongly supported if the triad is in root position or appears in a metrically strong position within a measure or larger musical unit.

The ratings given by Sandell are shown in Table 4.7 in rank order of their relative strengths. Although the values are based on a somewhat different set of considerations and exhibit a somewhat different use of the numerical rating scale, comparison with Table 4.6 shows good agreement between the two expert judges. For all but two measures (24 and 27), the judges agreed on the identity of the strongest key, and often on keys of lesser strengths. The second judge indicated an average of 3.29 different keys per measure, a value substantially greater than that found by the first judge. This difference is due in part to the second judge's criterion that the home key, C minor, is always present to some degree, a view held by Schenker (1906/1954).

Because both judges found that multiple keys were simultaneously present in many of the measures of the prelude, it is desirable to find a method for summarizing their judgments in a form that takes this feature into account. For this purpose, and for the purpose of comparing their judgments to the key-finding algorithm, we used the toroidal representation of musical keys shown in Figure 2.8. The objective is to find a point that best represents the relative weighting of the various keys in the experts' judgments and the algorithm's output.

First, however, some rather technical background is needed. Howard Kaplan (personal communication, November 1981) observed that if the toroidal multidimensional scaling solution is correct, the original probe tone rating profiles should contain two periodic components. This would give rise to the two circular projections found in the solution. Therefore, if a Fourier analysis were performed, then it should find two components with relatively large amplitudes. In particular, there should be one periodic component with five cycles per octave giving the horizontal dimension of Figure 2.8. This is because the circle of fifths is produced by taking the chroma circle (one cycle per octave) and wrapping it around five times. There should be another periodic component with three cycles per octave giving the vertical dimension of Figure 2.8. This dimension, which will be called the circle of thirds here, is such that keys the tonics of which

Table 4.7. Judgments of key strengths (Expert 2)[a] for Bach's C Minor Prelude (Book II)

Measure	Key 1	Key 2	Key 3	Key 4	Key 5
1	c 8.70	f 1.32			
2	c 8.70	f .66			
3	c 7.26	G 1.32			
4	c 6.66	B♭ 3.18	A♭ 1.32		
5	c 6.81	E♭ 2.58	A♭ 1.32	f .99	
6	c 5.16	E♭ 4.65	A♭ 1.86	f .66	
7	E♭ 7.26	c 5.16	A♭ 1.32	f .66	
8	E♭ 5.52	c 4.08			
9	E♭ 8.62	c 5.07	A♭ 3.57	B♭ .99	
10	E♭ 6.84	c 4.08	f .66		
11	E♭ 7.71	c 4.08	f .66	B♭ .66	
12	E♭ 8.26	c 5.07			
13	E♭ 9.75	c 4.08	B♭ .99		
14	E♭ 7.20	c 7.02	A♭ .99		
15	c 6.48	E♭ 3.21	A♭ 1.86		
16	f 5.52	A♭ 4.98	E♭ 2.22	c 1.50	D♭ .99
17	f 9.75	Eb 2.22	c 2.16	C .99	D♭ .66
18	f 5.10	E♭ 2.22	c 1.50	C 1.32	b♭ .66
19	f 5.65	A♭ 2.58	E♭ 2.22	c 1.50	b♭ .99
20	f 5.88	A♭ 4.98	c 1.50	E♭ 1.32	
21	f 7.74	b♭ 4.23	c 1.50		
22	f 7.54	c 1.50			
23	f 7.87	E♭ 4.44	c 1.50		
24	c 5.16	E♭ 4.65	f 2.22	A♭ 1.32	
25	c 9.96	f 2.22			
26	c 7.62	E♭ 4.44	f 2.22	A♭ .99	
27	f 5.76	E♭ 5.31	c 4.38	G 1.86	
28	c 10.00				

[a]Lower-case letters indicate minor keys.

are separated by a major third have the same vertical displacements, which results if the chroma circle is wrapped around three times. Thus, the two circular dimensions of the scaling solution should be reflected in the third and fifth harmonics of the Fourier expression. (It is entirely coincidental that the third and fifth harmonics correspond to the musical intervals called thirds and fifths; the harmonics refer to the number of periods per octave, whereas the musical intervals refer to the positions of the tones in the musical scale.)

Krumhansl (1982) obtained estimates of the constants in the Fourier expression, which is shown at the top of Table 4.8; the method used was that outlined in Jenkins and Watts (1968, pp. 17–21). Before computing the constants, the mean rating was subtracted from the data, so that the zeroth harmonic, which corresponds to the additive constant, had zero amplitude. The table shows the amplitudes, phases, and the proportion of variance accounted for by each of the other harmonics for the C major and C minor profiles. As expected, the third and fifth harmonics had the largest amplitudes for both profiles. The third harmonic was somewhat stronger for the minor key profile than the major key profile, whereas the fifth harmonic was stronger for the major-key profile than the minor-key profile. Profiles were resynthesized using only the third and fifth harmonics. These profiles correlated reasonably highly with the original data; the correlations were .91 and .89 for major and minor key profiles, respectively.

To substantiate the toroidal configuration, each of the 24 major and minor probe tone profiles was subjected to a Fourier analysis. This made it possible to generate a spatial representation of the full set of keys from the estimated phases. The spatial representation, with two angular dimensions corresponding to the phases of the third and fifth harmonics, was virtually identical to the scaling solution. Given that only the third and fifth harmonics were used, the configuration determined by the phases necessarily contained the circle of thirds and the circle of fifths, but it also placed the major and minor keys in the same relative positions as in the multidimensional scaling solution.

For the present application, Fourier analysis provided an efficient method for determining points in the toroidal configuration, where the points reflect the relative weights of different keys. Another method for accomplishing the same objective, multidimensional unfolding, was used by Krumhansl and Kessler (1982) (see Chapter 9), but it was not employed in this case because of its greater computational demands. (In

Table 4.8. Fourier analysis of major and minor profiles

$$f(x) = \sum_{m=1}^{6} \mathbf{R}_m \cos(mx - \phi_m)$$

	Major Profile			Minor Profile		
Harmonic	$\mathbf{R}_m$	ϕ_m	Percentage of Variance	$\mathbf{R}_m$	ϕ_m	Percentage of Variance
1	.051	242.9	.003	.166	336.4	.041
2	.318	184.6	.127	.191	120.0	.055
3	.404	5.0	.205	.469	51.4	.330
4	.147	142.2	.027	.270	192.8	.109
5	.711	35.5	.634	.557	334.5	.465
6	.086	180.0	.005	.009	180.0	.000

cases when we have applied the two methods to the same data, we have found similar results.)

The following procedure was used to obtain a point for each measure of the prelude for each of the two judges. For each key receiving a nonzero value, the probe tone rating profile was multiplied by the judges' estimates of key strengths and the sum of the resulting vectors was computed. For example, the first judge gave for measure number 4 a value of 9.42 to C minor, so the 12-dimensional vector for C minor (the probe tone ratings) was multiplied by this value. To this vector was added the vector obtained by multiplying the 12-dimensional vector for B♭ major (its probe tone ratings) by the value 3.81. The weighted sum of the vectors was then subjected to a Fourier analysis and the phases found for the fifth and third harmonics were taken as the coordinates in the toroidal configuration. For this example, the procedure produced a point between C minor and B♭ major. A point was found in this way for each measure for each of the two judges; they are plotted in Figures 4.7–4.12 as the dashed lines.

We tried various schemes for modeling the experts' judgments using the distributions of tone durations in the piece. In one, we used the distributions for each measure separately; this distribution was then subjected to Fourier analysis to obtain a point in the torus representation.

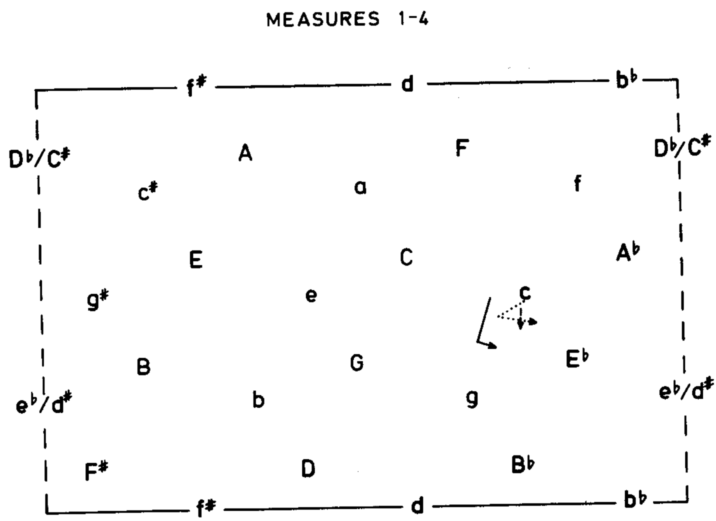

Fig. 4.7. The points obtained from the experts' judgments of key strengths (broken lines) and the model's analysis of tone durations (solid line) for measures 1–4 of Bach's C Minor Prelude, Book II. The model uses the tone durations in each measure weighted twice and the tone durations in preceding and subsequent measures each weighted once.

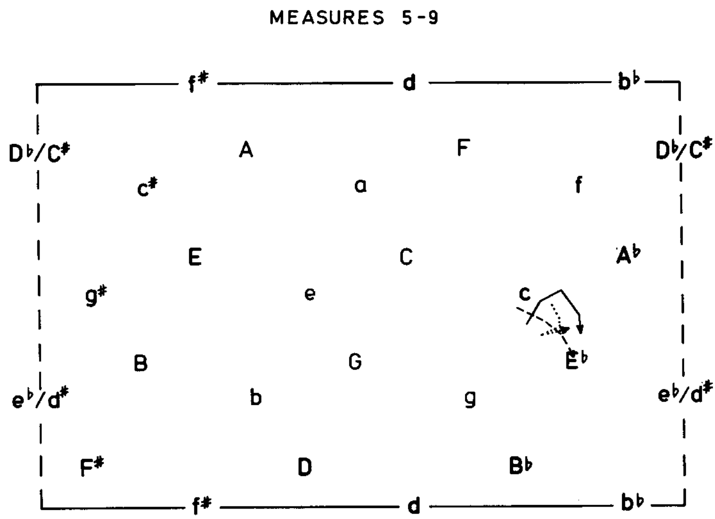

Fig. 4.8. The points obtained from the experts' judgments of key strengths (broken lines) and the model's analysis of tone durations (solid line) for measures 5–9 of Bach's C Minor Prelude, Book II.

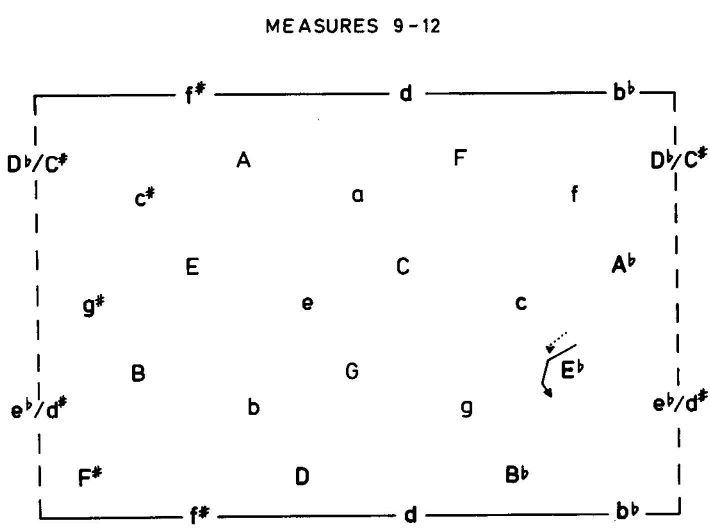

Fig. 4.9. The points obtained from the experts' judgments of key strengths (broken lines) and the model's analysis of tone durations (solid line) for measures 9–12 of Bach's C Minor Prelude, Book II.

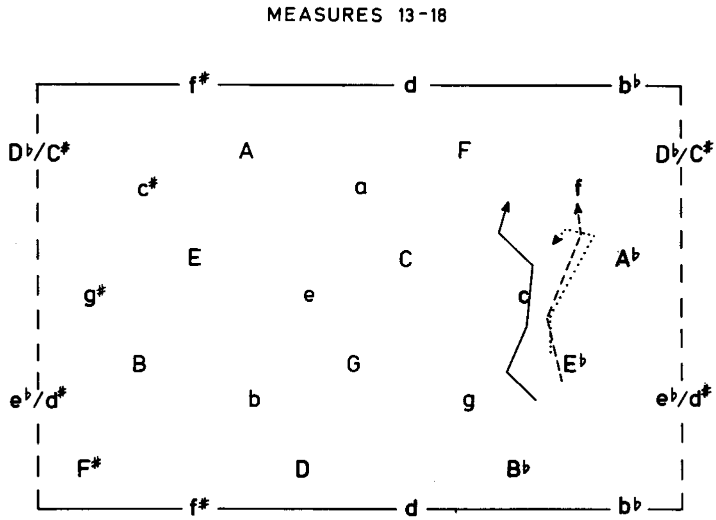

Fig. 4.10. The points obtained from the experts' judgments of key strengths (broken lines) and the model's analysis of tone durations (solid line) for measures 13–18 of Bach's C Minor Prelude, Book II.

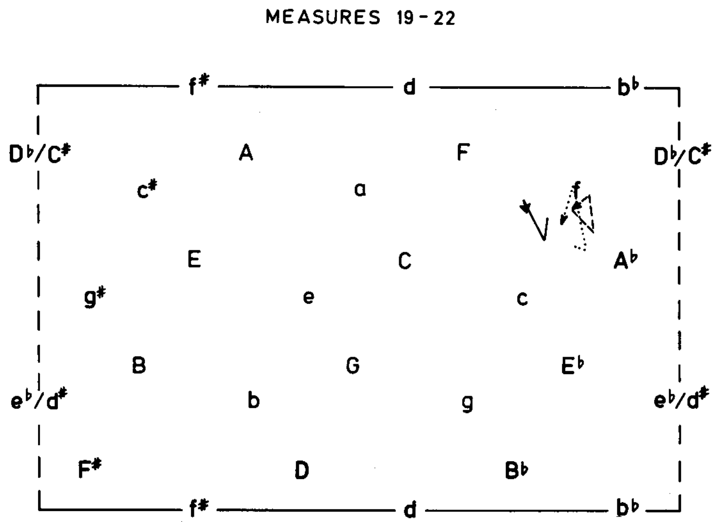

Fig. 4.11. The points obtained from the experts' judgments of key strengths (broken lines) and the model's analysis of tone durations (solid line) for measures 19–22 of Bach's C Minor Prelude, Book II.

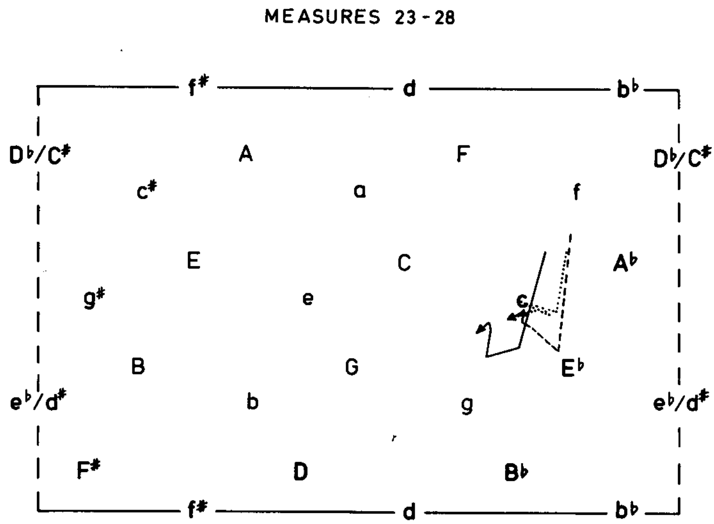

Fig. 4.12. The points obtained from the experts' judgments of key strengths (broken lines) and the model's analysis of tone durations (solid line) for measures 23–28 of Bach's C Minor Prelude, Book II.

Because this did not provide a completely satisfactory fit to the experts' judgments, we used various weighted sums of the duration distributions that included preceding and/or subsequent measures of the prelude. In no case did the input vector exceed three measures. The weighted sums of the duration distributions were then subjected to Fourier analysis. Of the various weighted models tried, the best match to the experts' judgments was obtained when the preceding measure had a weight of one, the current measure had a weight of two, and the subsequent measure had a weight of one. That this model provided a better fit to the experts' judgments than the other models suggests that their ratings of key strengths were based on information both preceding and following the measure in question. The points obtained using this model are shown in Figures 4.7–4.12 as the solid line.

In general, the weighted model agreed well with the experts' judgments. The first phrase of the prelude, measures 1–4 corresponds to points in the region of C minor (Fig. 4.7). The second phrase, measures 5–9, shows a shift from the C minor to the E♭ major region (Figure 4.8), where the points remain until the end of the next phrase in measure 12 (Figure 4.9). The next phrase, in measures 13–18, contains a pattern of shifting tonal orientations moving from E♭ major to the region around F minor (Figure 4.10). The following phrase, measures 19–22, remains in the F minor region (Figure 4.11), followed by a shift back to the original key of C minor in the final phrase in measures 23–28 (Figure 4.12). This pattern is apparent not only in the points for the two experts, but is also followed quite closely by the points found by the Fourier analysis of the weighted duration distributions.

To quantify the discrepancies between the three sets of points (for the model and the two experts), the square root of the sum of the deviations squared was computed for each measure. This corresponds to the distance between the points in the rectangular configuration shown in Figures 4.7–4.12. For the two experts, this value averaged 14.3 degrees. For the model and the first expert, this value averaged 23.3 degrees; for the model and the second expert, this value averaged 25.0 degrees. Therefore, there was somewhat greater agreement between the two experts than between the model and either one. For the unweighted model using only the duration distributions for each measure separately, the deviations of the model from the two experts averaged 33.6 and 35.5, respectively. Therefore, the weighted model clearly provided an improvement over the single-measure model.

A better fit to the experts' judgments might have been obtained if the analysis were based on longer segments of the prelude, rather than on input samples consisting of a maximum of three measures as used here. However, this probably would have resulted in a loss of sensitivity to the shifting tonal centers contained in this prelude, which is modeled quite well in this application. Clearly, the computer analysis lacks some of the precision found in the experts' judgments, which are based on examination of specific chromatic and harmonic features of the music. However, the extent to which the model is successful suggests, at a minimum, that it is able to identify those sections containing tonal ambiguity and modulations, relatively subtle features of the music, found by our expert judges.

Limitations and possible extensions of the algorithm

The present algorithm uses input vectors containing only information about tone durations, which are matched to the tonal hierarchies of major and minor keys. This is not to imply, however, that other features could not be used instead or in addition. The success of the algorithm demonstrates that this information would be sufficient in many cases to determine the correct key region, if not precisely the correct key. Other kinds of information might be used in conjunction with the tone durations to sharpen the key judgments. Various possibilities include the order of tones, melodic patterns, chords and chord sequences, and points of metrical and rhythmic stress. In what follows, we make some suggestions concerning the potential role of these variables and how they might be incorporated into a computer algorithm.

The importance of tone order was demonstrated by Krumhansl (1979) in both relatedness judgments of pairs of tones and memory confusions (this and related studies are described in the next chapter). These studies found that, for two tones differing in structural significance in a tonal context, the less stable tone was judged as more related to the more stable tone than the reverse. For example, the ordered pair of tones B C received a higher rating in a C major context than did the ordered pair C B. Similarly, less stable tones were confused relatively often with more stable tones compared to the reverse order. Moreover, the relatedness of tones depends on context. Although B and C, for example, are perceived as quite strongly related in C major, they are not perceived as strongly related in B major. These effects are also reflected in statistics of two-tone sequences in tonal melodies (Youngblood, 1958; see also Table 5.3).

This suggests it would be possible to supplement the present algorithm by a matrix specifying the perceived degree of relatedness between successive tones in each of the 24 keys. For a particular input segment, the average of the ratings for successive tones would reflect the likelihood that the sequence would appear in each possible key context. As will be seen, one influence determining perceived relations between tones is their position in the tonal hierarchy. Therefore, the results based on relatedness judgments and the tonal hierarchy would be expected to converge on the same key. The two sources of information in combination, however, would probably sharpen the key-finding process. These intuitions were supported by initial explorations in which we augmented the basic algorithm described here with perceptual judgments of melodic intervals in tonal contexts. We decided not to include order information in the algorithm described here primarily because the short input segments used in the applications contain many simultaneous tones, making the coding of order difficult; order is coded easily only for single-voiced melodies. In addition, an empirical assessment is presently lacking of the relative perceptual salience of the tonal hierarchies and order information, which would be needed in such an extension.

The series of experiments reported by Bharucha (1984a) also emphasizes the importance of order information. Briefly, his principle of melodic anchoring states that a tone will become anchored to a following tone if the first approaches the second by stepwise motion along a chromatic or diatonic scale. Under these conditions, the first tone becomes assimilated to the tonal framework of the second tone or, in other words, the second tone contributes more to the sense of key. The melodic-anchoring principle, supported empirically by experiments using a variety of tasks and stimulus materials, could be implemented in a key-finding algorithm in the following way. Each sequence of tones could be evaluated according to whether it satisfies the conditions of melodic anchoring; if so, greater weight would be given to the tones that serve as anchors. Additional experimentation and modeling would be needed to assign appropriate weights and determine how these interact with other variables such as tone duration and rhythm. In addition, plausible diatonic scales need to be determined before evaluating whether a sequence of tones satisfies melodic anchoring. The present algorithm might be used for this purpose. We suggest, then, that the two approaches might usefully be combined.

Brown and Butler (1981; Butler, 1982; Butler & Brown, 1984; Brown, 1988) have also stressed the importance of order information. Their focus, however, has been on the "rare" intervals, minor second and tritone (Browne, 1981). These are called rare because the minor second appears just twice in the diatonic scale (between the third and fourth scale degrees, and between the seventh scale degree and the tonic), and the tritone appears just once (between the fourth and seventh scale degrees). Because of this, these intervals are distinctive features of each key and could serve as important cues for finding the tonality, a hypothesis supported by their experimental results. This suggests that a key-finding algorithm might be devised that gives special weight to the rare intervals. In conjunction with other distributional information, this might yield highly accurate results.

Another source of information that is clearly important for determining key is harmonic structure, particularly the presence of structurally significant chords and conventional chord cadences. Winograd's artificial-intelligence model (1968) probably remains the single most comprehensive and successful contribution to the computer analysis of tonal harmony. At each point in the parsing process, each chord (or group of chords) is described by its function within the nested structure of tonalities in which it operates. However, because of the inherent ambiguity of harmonic functions, multiple parsings are possible. To deal with this problem, Winograd invokes two "semantic" measures: a tonality hierarchy and a system of plausibility values for selecting among possible parsings. The first is a measure of the likelihood of various possible keys, and the algorithm proposed here might provide a simple method for determining these values. Other information, possibly derived from perceptual judgments of chords (e.g., Krumhansl, Bharucha, & Kessler, 1982; Bharucha & Krumhansl, 1983; Krumhansl, Bharucha, & Castellano, 1982; these studies are reviewed in a later chapter) or statistical summaries of harmonic music, might then be employed as plausibility values to guide the search for the most likely parsing. The extent to which the guided search finds expected progressions, cadential formulas, and other harmonic conventions might in turn provide feedback to sharpen the key assignments in an interactive fashion. Thus, in this kind of artificial-intelligence context, the present algorithm might serve as a preprocessing stage for more detailed computer analysis of tonal music.

A final source of information that may contribute to the sense of key is rhythmic or metrical structure. In general, one would expect structurally significant pitch events to fall at points of stress or accent. This suggests that assigning greater weight to these events would improve the key-finding algorithm. The choice of stress values might be guided by theoretical treatments (e.g., Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983; Benjamin, 1984; Longuet-Higgins & Lee, 1984) or empirical measures. The studies by Palmer and Krumhansl (1987a,b) introduced two methodologies for assessing rhythmic structure in a form that might be useful in such an algorithm (these studies are described briefly in Chapter 6).

For several reasons, this approach might encounter difficulties, however. At least, attempts to incorporate rhythm into a model such as that proposed here may be premature. Many theoretical treatments of rhythm are based, at least in part, on pitch information and more particularly on the structural significance of pitch events (e.g., as in harmonic rhythm), which requires prior knowledge of the tonality. Second, theoretical treatments presently lack detailed empirical evaluation and quantification. An exception to this is the partial confirmation found by Palmer and Krumhansl (1987a,b) of the metrical stress hierarchies of the theory of Lerdahl and Jackendoff (1983). What is more important for the present discussion is that these studies found that judgments of temporal patterns were independent from judgments of pitch patterns. For the particular musical selections employed, tonally strong events did not consistently occur at points of rhythmic stress.

The possibility of this kind of independence between pitch and rhythmic organizations was also suggested by a study of Knopoff and Hutchinson (1981). In a series of analysis of 14 Bach chorales (with a total of 1263 chords), chords on strong beats were not more consonant than chords on weak beats. Presumably, consonance correlates with stability or finality; therefore this again suggests that pitch and metrical patterns may be independent. A final reason for supposing that rhythmic stress is not necessary for determining key is that A. J. Cohen's listeners (1977) were quite accurate when they heard segments so brief that it is unlikely that a rhythmic organization would have had time to become established. These considerations suggest that further investigation of the relationship between rhythmic and pitch structures is needed before attempting to introduce rhythm into models of key-finding.

To conclude, the algorithm suggests one kind of process that may contribute to the listener's sense of key. The sounded events, matched against internalized tonal hierarchies, could yield useful hypotheses about possible keys. This could then set up a framework that would serve to guide expectations about ordered events, melodic patterns, harmonies, and metrical structure and phrasing. These other variables, in turn, may reinforce or contradict the sense of key to varying degrees. Some combination of variables and their interactions would be expected to yield higher accuracy rates than the present algorithm. However, it is important to bear in mind that music is, in some cases, intentionally structured to place various structural properties in opposition, creating perceptual ambiguity and tension that is resolved only when the separate components finally come into correspondence.