Music Perception
Summer 2000, Vol. 17, No. 4, 461–479

Tonality Induction: A Statistical Approach Applied Cross-Culturally

CAROL L. KRUMHANSL
Cornell University

Sensitivity to tone distributions has been proposed as a mechanism underlying tonality induction. This sensitivity is considered in a cross-cultural context using two styles of music, Finnish spiritual folk hymns and North Sami yoiks. Previous research on melodic continuation judgments showed strong correlations with the statistics of the musical style, specifically, the tone distributions and two- and three-tone transitions. This article develops models using these three kinds of statistics to categorize short initial segments as coming from one style or the other. The model using tone distributions was found to make numerous categorization errors, which can be understood because the tone distributions for these styles are similar. However, categorization was better for the models that used two- and three-tone transitions. The major differences between the transitional probabilities in the styles were analyzed, and these differences were used to account for the cases that the models found difficult. These results point to listeners' sensitivity to higher order transition information and its utility for style identification.

THIS article explores in a cross-cultural setting basic psychological principles that may underlie tonality induction. In tonal-harmonic Western music, tonality induction is the process through which a listener identifies the key of a piece of music, that is, what the tonic is and whether the key is major or minor. Identifying the tonality is an important first step in analyzing a piece of music. Once the key is known, then the piece can be analyzed harmonically, its grouping and phrase structures can be determined, its formal organization can be specified, and so on. In other words, tonality (together with meter) usually precedes further theoretical analysis of the music. From a psychological point of view, identifying the tonality is important for encoding and remembering the music. Extensive experimental research (summarized, e.g., in Dowling & Harwood, 1986; Handel,

Address correspondence to Carol L. Krumhansl, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853. (e-mail: clk4@cornell.edu)
ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

1989; Krumhansl, 1990) has shown various effects of tonality. Among other effects, tones in the key are better remembered, are identified more rapidly, are preferred in final temporal positions, contribute to phrase ending, are more expected than nonscale tones, and are less frequently confused with nonscale tones than the reverse. More generally, tonality induction is interesting psychologically because it appears to rely on two basic principles of perception and cognition.

The first principle is the existence of cognitive reference points (Rosch, 1979). Cognitive reference points refer to instances within categories that are used to encode, name, and remember other instances within the category. For example, in the category of fruit, some elements, such as apple, function as cognitive reference points. Other instances are described in reference to these, for example, in the names given to pineapple and *Apfelsine* ("apples from China" in German). Cognitive reference points are held to be an efficient mechanism for categorization, requiring a minimum memory load and generally enabling instances to be correctly categorized. In music, the tonality establishes certain tones as reference points, creating a hierarchy of tones. The tonic heads the hierarchy, followed by the fifth and third scale degrees, followed by the other scale tones, and finally the nonscale tones. One method used to study this experimentally is the probe-tone technique (Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979). Musical contexts (e.g., scales, tonic triads, chord cadences, or short excerpts from pieces) are followed by a single tone, called the probe tone. Instructions variously ask how well the tone completes or fits with the context.

The second principle is sensitivity to the frequencies with which instances occur. Psychological research has shown quite accurate memory for how often different instances have been presented during the learning phase of memory experiments. Various results also show that instances that occur more frequently are more easily encoded, named, and remembered. Frequent instances are also considered to be better category members. For example, Rosch (1979) found that the word "apple" occurred more frequently in a corpus of written English than did the word "pineapple." This suggests links between cognitive reference points and sensitivity to frequency. Cognitive reference points may be established initially by their relatively frequent occurrence. In music, tones given higher ratings in the probe-tone task are also sounded more frequently in Western music (Krumhansl, 1990). This suggests that the reference points are established initially by frequent repetition. Supporting evidence comes from a study that used North Indian classical music (Castellano, Bharucha, & Krumhansl, 1984). In that study, Western listeners, who were unfamiliar with the style, gave responses similar to the responses of Indian listeners. This similarity was attributed to Western listeners' basing their responses on the distributions of tones in the contexts presented, that is, the frequency and duration with which the tones were sounded. Frequency effects were also found by Oram and Cuddy (1995), who constructed novel tone sequences according to predetermined (and novel) frequency distributions. They showed that these novel frequency distributions induced a tonal hierarchy, as measured in the probe-tone task. Frequency sensitivity may play an especially central role in music, because most listeners do not have absolute pitch and rely primarily on relative pitch. The frequent repetition of the reference tones provides a framework for encoding and remembering the sounded tones.

Tone Frequencies and Tonal Hierarchies in Tonality Induction

The research just summarized shows a connection between cognitive reference points in music and distributions of tones. This apparent connection led to the development of a key-finding algorithm (by myself and Mark Schmuckler and described in Krumhansl, 1990). The algorithm was proposed as a possible model of how a listener might identify the key initially and track modulations from one key to another. One premise of the model is that listeners have internalized tone frequencies in pieces in major and minor keys through prior experience. The second premise is that when listeners hear a section of a piece of music, they match the distribution of sounded tones to the internalized tonal hierarchies. The algorithm begins with a sample of tones, for example, each measure of a musical piece. The total duration (or number of occurrences) of each of the chromatic scale tones is then determined. This tone distribution is then correlated with the tonal hierarchies of all major and minor keys (using the probe-tone values from Krumhansl & Kessler, 1982). The highest correlation is taken to be the key identified by the algorithm.

The first application of the algorithm was to the initial four-tone segments of the piano preludes of J. S. Bach, Shostakovich, and Chopin. For the 48 Bach preludes, the algorithm identified the correct (composer's designated) key 44 times with just the first four tones. When an incorrect key was found, it was closely related to the correct key. For the 24 Shostakovich preludes, the algorithm was correct on 17 preludes, and a closely related key was identified for the other preludes. The algorithm was far less accurate on the Chopin preludes, identifying only 11 keys correctly on the basis of the first four tones. Of the remaining 13 preludes, 8 Prelude excerpts produced closely related keys. In three more cases, a diffuse collection of keys near the correct one was identified, with no probe-tone profile correlating strongly with the tone distributions. In the last two cases, many other keys had higher correlations than the correct key had. Both these prelude segments contained a chromatic alteration of one of the scale tones (the tonic in one case, and the third scale degree in the other). Thus, it is not surprising that the algorithm failed in these cases to identify the key from the four-tone segments.

The second application was to the fugue subjects of J. S. Bach and Shostakovich. Here, the question was how many tones would be needed to find the correct key. In the case of the Bach preludes, the correct key was found 44 times out of 48, with an average of 5.1 tones needed for the correct identification. In the case of the Shostakovich preludes, the correct key was found 22 out of 24 times, with an average of 4.3 tones needed for the correct identification. Both these applications suggest that the short initial portions of these pieces tend to emphasize, by repetition and duration, those tones that are relatively high in the tonal hierarchy. Although errors occurred, the correct identifications were far more frequent than one would expect by chance (1 out of 24). Most of the errors were identifying closely related keys, allowing a process such as that modeled by the algorithm to orient the listener to the correct tonal region, if not the correct key.

The third application was to a single prelude (C Minor, Book II) of J. S. Bach. We examined the entire piece on a measure-by-measure basis. The prelude contains an interesting pattern of modulations, and the question was whether the algorithm could track these modulations. In this case, we also had experts' judgments; they were asked to indicate the primary key for each measure and also any keys of lesser strength. The experts appeared unable to focus exclusively on single measures, showing influences of preceding and following measures. Thus, the algorithm's input consisted of the tone durations in the current measure (with full weights) and the tone durations in the preceding measure and the following measure (each with one-half weights). The algorithm showed good general agreement with the experts, although it lacked some of the precision found in the experts' judgments. However, the degree to which the algorithm was successful suggests that it is able to identify sections containing tonal ambiguity and modulations in general agreement with the experts.

The Krumhansl-Schmuckler algorithm is simple, using only prior psychological data and the durations of tones in short excerpts. It does not carry out any music-theoretical procedures, such as harmonic analysis (see, e.g., Rowe, 1993, 2000; Vos, 2000; Vos & Van Geenen, 1996). Nor does it take meter into account, or phrasing. The most important missing element for present purposes, however, is that the algorithm does not take the temporal order of the tones into account. Initial explorations, described briefly in Krumhansl (1990), suggested that the algorithm could be sharpened by the addition of information about tone order. In the extension of the algorithm, the number of two-tone transitions (how often one tone was followed by another) was compared with listeners' judgments of melodic intervals in tonal contexts (Krumhansl, 1979, 1990). This method, however, could be applied readily only to monophonic music because order is coded easily only for single-voiced melodies. Thus, we did not develop it further, although extensive psychological data show the importance of tone order as is discussed next.

The Importance of Temporal Order in Music

In tonal contexts, the order of tones is highly important. This kind of order dependency reflects the temporal nature of music, with auditory events occurring along the asymmetric dimension of time going forward in one direction. Temporal-order effects in music have been treated experimentally at least from the beginning of the 20th century (see Krumhansl, 1990, p. 122 for citations). These early studies demonstrated differences in judgments (such as preference and finality) as a function of the order in which the tones were heard. A scaling study (Krumhansl, 1979) studied temporal-order asymmetries of all pairs of tones in an octave range. Listeners were asked to make similarity judgments on sequentially presented pairs of tones in a clear tonal context. Specifically, following the tonal context, two tones were heard and the listeners were asked to judge how closely related the first tone is to the second tone in the tonal system suggested by the context. The ratings for pairs of tones were often asymmetric, that is, they depended on the order in which the two tones were presented. The asymmetries followed a regular pattern. Higher similarity ratings were given when the first tone was lower in the tonal hierarchy than the second. In other words, two tones were judged as more similar when the second tone was relatively high in the tonal hierarchy, such as the tonic, dominant, and third.

Two other experiments (Krumhansl, 1979) found temporal-order effects in memory confusions. A to-be-remembered tone was sounded before a melodic sequence. After the sequence, another tone was sounded and the listeners had to judge that tone as the same as or different from the to-be-remembered tone. In both experiments, listeners confused the previously heard nonscale tones more frequently with scale tones than they confused the previously heard scale tones with nonscale tones. Thus, both similarity and memory judgments contained a regular pattern of temporal-order asymmetries, leading to the conclusion, "These results suggest that, in the psychological representation, those tones less closely related to the tonality are less stable than tones closely related to the tonality, and that the representation incorporates the tendency for less stable tones to move toward more stable tones in time. Thus, these temporal asymmetries reflect the dynamic character of musical tones in time" (Krumhansl, 1979, p. 372).

Similar temporal-order effects were found in various studies done in collaboration with Jamshed Bharucha (summarized in Krumhansl, 1990) for chord sequences. The effects were large and appeared in both similarity and memory judgments. They could be manipulated in a predictable and interpretable way by varying the tonal context. The temporal-order effects for both tones and chords could be related to tonal and harmonic hierarchies. And, most important in the present context, the psychological judgments correlate with the frequency of chords and chord progressions in music. In sum, it appears that frequency of successive elements in music is incorporated in listeners' cognitive representations and that this information may be used to orient to the tonality of a piece of music. This possibility will now be considered in an algorithm that uses tone-order information. The algorithm will be applied to two styles of vocal monophonic music, Finnish spiritual folk hymns and North Sami yoiks, that were used in two recent studies.

Two Cross-Cultural Studies: Finnish Spiritual Folk Hymns and North Sami Yoiks

The two studies investigated melodic expectancy. Expectancy has an important function in a wide variety of behaviors, including perception, speech understanding and production, and skilled performance. Diverse methods have been developed to study musical expectancy, including production, memory, detection, priming, and structural judgments. The two studies grew out of a collaboration between myself, Jukka Louhivuori, Petri Toiviainen, Topi Järvinen, Pekka Toivanen, Tuomas Eerola, all from the University of Jyväskylä, and Annukka Hirvasvuopio from Utsjoki and the University of Tampere (Krumhansl, Louhivuori, Toiviainen, Järvinen, & Eerola, 1999; Krumhansl, Toivanen, Eerola, Toiviainen, Järvinen, & Louhivuori, 2000). Our approach was to compare results from behavioral experiments, a statistical analysis of the musical styles, and a neural network model of the self-organizing type (Kohonen, 1997).

In both behavioral experiments, eight test excerpts were chosen from each of the two styles of music, Finnish spiritual folk hymns and North Sami yoiks. They were chosen so that they ended when expectations for continuation would be expected to be strong, and the actual music contained an unexpected continuation that might be known to the more expert listeners but not to the others. The excerpts were transposed so that they fell within the range from G_3 (the G below middle C) to E_5 (the E an octave and a third above middle C) and had C as a reference tone (tonic in Western music). The method used in the behavioral experiment was similar to that used in earlier research on musical expectancy (e.g., Krumhansl, 1995; Schmuckler, 1989). Each trial began with the excerpt followed by a single tone, called the continuation tone. All continuation tones within the range (from the G_3 to the high E_5) were used on different trials in a random order. The listeners were asked to judge how well the continuation tone fits with their expectations about what might follow in the melody.

The Finnish spiritual folk hymns combine elements of Lutheran hymns and Finnish folk songs. The style is generally associated with the Beseecher religious prayer movement, which created the texts and melodies in the middle of the 18th century in southwest Finland. Two examples are shown in Figure 1. For the study on Finnish spiritual folk hymns, two groups of listeners participated. The first group consisted of 12 subjects who were members of the Youth Choir of the Beseechers. The experiment was run at their rehearsal in the parish hall of Eura, Finland. They had heard and learned the spiritual folk hymns at an early age and were highly familiar with the style, and most subjects knew most of the hymns in the experiment. The second group consisted of 14 undergraduate musicology majors at the University of Jyväskylä. They were not familiar with the particular hymns and did not recognize any of the melodies. However, they reported being familiar with the style that is to be expected because the spiritual folk hymns combine elements from Lutheran hymns and Finnish folk music.

The Sami yoiks, from Lapland, have a style quite distinct from Western music. Yoiks are almost exclusively vocal and typically use a few nonsense syllables rather than words. Yoiks mostly deal with topics of everyday life,

HSHL 2a (Number 1)

HSHL 20 (Number 4)

Fig. 1. Two Finnish spiritual folk hymns transposed to C major.

such as people, animals, places, and nature. They have personal meaning for the yoiker and are sung either in solitude or with people who share knowledge of the yoik's topic. Yoiks are highly improvisatory and have not been extensively analyzed from a theoretical point of view. They are characterized by short, repeated motives that often comprise only a few notes, in most cases from the major pentatonic scale. They tend to include relatively large interval leaps. Examples are shown in Figure 2. Three groups of listeners participated in the study on North Sami yoiks. The first consisted of seven native Sami who had extensive experience learning and performing yoiks. They reported that they were highly familiar with the style, although they were very familiar with only two of the yoiks selected as experimental materials. The second group consisted of 11 undergraduate music education majors at the University of Jyväskylä. Their primary training was in Western music. However, they had learned to play a number of yoiks during a 1-year period, including two of the yoiks used in the experiment. The third group consisted of Western musicians unfamiliar with the yoiks and the yoik style. This allowed us to separate two factors: (1) musical culture and (2) experience with the particular yoiks.

The following summary focuses exclusively on the convergence between the melodic continuation judgments and the statistical style analysis. The statistical style analysis considered a corpus of 18 Finnish spiritual folk hymns (only the 9 hymns in major will be considered here) and 18 yoiks. For each corpus, the analysis found the tone distributions (how frequently

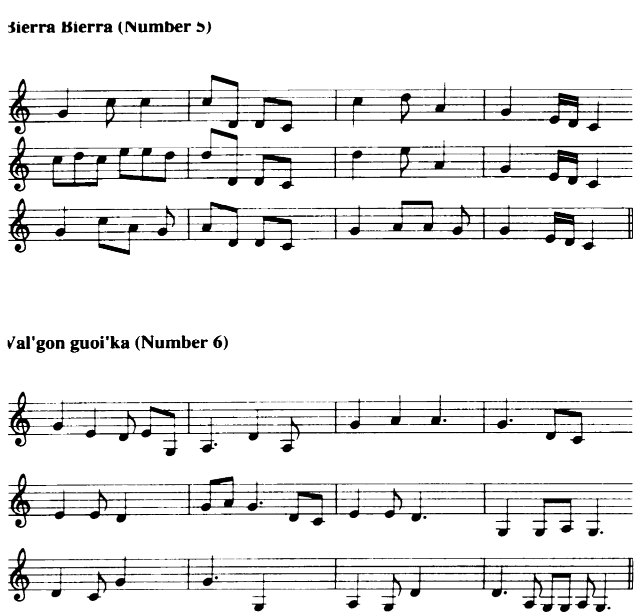

Fig. 2. Two North Sami yoiks transposed so that C and G are frequent tones.

each chromatic scale tone appeared) and the frequency of all possible two- and three-tone transitions.

The correlations with the behavioral results from the Finnish spiritual folk hymn study (Krumhansl et al., 1999) are shown at the top of Table 1. As seen in the first line, the melodic continuation judgments of both groups of listeners correlated strongly with the distributions of tones in the corpus. (The degrees of freedom of these and all the other correlations in the table is 174.) The next line shows the correlations between the judgments and the two-tone transitions, that is, the relative probabilities of the continuation tone conditional on the last tone of the experimental contexts. Again, these correlations were highly significant. The next line shows the correlation between the judgments and the three-tone transitions, that is, the relative probabilities of the continuation tone conditional on the last two tones of the experimental contexts. Again, these correlations were highly significant, and here a difference between the groups was found such that the correlations for the experts were reliably higher than the correlations for the nonexperts. Finally, the correlations with a variable coding the cor-

TABLE 1
Correlations Between Statistics of Music and Behavioral Judgments of Melodic Continuations

	Spiritual Folk Hymns		
	Subject Group		
	Finnish	**Beseecher Choir**	
Tone distributions	.64***	.72***	
Two-tone transitions	.50***	.56***	
Three-tone transitions	.41***	.59***	
Correct next tone	.15	.52***	
	North Sami Yoiks		
	Subject Group		
	Western	**Finnish**	**Sami**
		All Yoiks	
Tone distributions	.63***	.74***	.66***
Two-tone transitions	.67***	.61***	.57***
Three-tone transitions	.41***	.38***	.43***
Correct next tone	.11	.21**	.32***
		Yoiks Known Only to Sami	
Tone distributions	.65***	.78***	.67***
Two-tone transitions	.69***	.60***	.54***
Three-tone transitions	.49***	.46**	.65***
Correct next tone	.07	.18	.50***

$^{**}p < .01$, $^{***}p < .001$.

rect next tone of the hymn was much stronger for the experts than for the nonexperts. In terms of Bharucha's (1987) distinction between schematic and veridical expectations, the experts showed stronger veridical expectations than the nonexperts.

The correlations with the behavioral results from the Sami yoik study (Krumhansl et al., 2000) are shown in the rest of Table 1. Again highly significant correlations were found between the continuation tones and tone distributions, the two-tone transitions, and the three-tone transitions. The correlation with the correct next tone (veridical expectations) was strongest for the Sami, followed by the Finnish, and finally the Western subjects. Recall that only some of the yoiks were familiar to the Sami subjects. The results for this subset of yoiks are shown at the bottom of Table 1. Here the effect of veridical expectations is more pronounced. In addition, for these yoiks, the correlation with the higher order three-tone transitions was noticeably higher for the Sami than for the other two groups of listeners. The central point to take from both studies is that listeners, even those with little prior familiarity with the style, appear to be sensitive to statistical distributions of tones. This result is similar to those of the North Indian music study (Castellano et al., 1984). The present results show that listeners are not only sensitive to the distributions of tones, but also to higher order statistics, such as the two- and three-tone transitions.

Statistical Distributions and Musical Style

These findings stimulated further analysis of the problem of tonality induction. They suggested that it might be interesting to develop an extension of the Krumhansl-Schmuckler algorithm (Krumhansl, 1990) that used not only the tonal hierarchies (which correlate with the frequencies of single tones) as was done in the earlier algorithm, but also the frequencies of successive tone combinations. In addition to evaluating the utility of this extension of the earlier algorithm, this approach may be especially useful in a cross-cultural context where the music has not been treated theoretically. Using the statistics of the music may yield insights into how a theoretical analysis of the style might be developed. The two styles of music used in these studies, Finnish spiritual folk hymns and North Sami yoiks, lend themselves to this analysis because music in both traditions is monophonic, thus eliminating the problem of how to code relationships between successive tones. The particular focus in what follows is on whether the algorithm can correctly distinguish between hymns and yoiks given the initial seven tones.

Figure 3 shows the tone distributions from the corpus of 9 hymns in major and 18 yoiks. Both sets of values were normalized before being

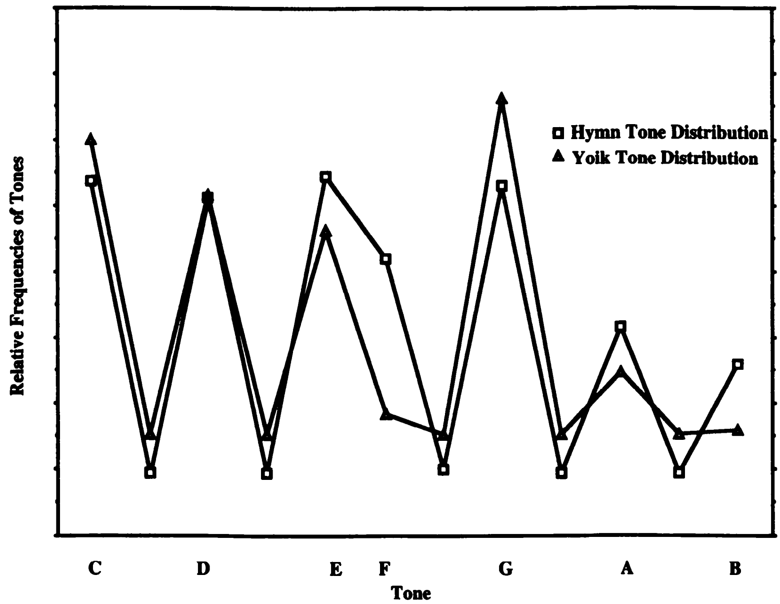

Fig. 3. Tone distributions for a corpus of 9 Finnish spiritual folk hymns transposed to C major and a corpus of 18 North Sami yoiks transposed so that C and G are frequent tones.

graphed. Two points should be noted. First, the tone distributions for the major hymns and yoiks were very similar (with a correlation of r (10) = .87, $p < .0005$). This correlation is higher than the correlation between the tonal hierarchies of any major or minor keys (Krumhansl, 1990, p. 38). Thus, the problem of distinguishing between these styles on the basis of the tone distributions would be harder than distinguishing between any two major and minor keys on the basis of their tonal hierarchies. Second, although the hymn and yoik tone distributions are very similar, they differ in minor ways. They differ in as much as the F and B (and to some extent A) appear relatively infrequently in the yoiks compared with the hymns. Also, the G appears more frequently than C in the yoiks, whereas this difference does not appear in the hymns. For this reason, the yoik hierarchy might be said to be pentatonic in orientation and to have both C and G as reference points.

Three models were tested with the materials listed in Table 2. All eight Finnish spiritual folk hymns used to test the model were transposed to C major and began with the tones shown in the table. The eight yoiks were those used in the behavioral experiment, which were transposed so that C and G were relatively frequent. Their initial seven tones also are listed in Table 2. Seven tones were chosen for the tests because the previous application (Krumhansl, 1990) found this was approximately the number of tones

TABLE 2
Materials to Test the Models

Finnish Spiritual Folk Hymns		
Hymn No.	**Halullisten Sielujen Hengelliset Laulut (HSHL)**	**First Seven Tones**
1	HSHL 2a	CCBABCD
2	HSHL 54a	CCEGGAB
3	HSHL 12a	GGECEGG
4	HSHL 20	EGCCBAD
5	HSHL 11a	CDEFDGA
6	HSHL 15a	CDEFDGG
7	HSHL 17	GCCCDED
8	HSHL 32	CBCDEDC

North Sami Yoiks		
Yoik No.	**Name of Yoik**	**First Seven Tones**
1	Anden Inga	CCCDDCD
2	Elle Sunna	GCDDCGD
3	¢appa Magdalena	CDGCDCE
4	Haldí	EEEDCDD
5	Bierra Bierra	GCCCDDC
6	Val´gon guoi´ka	GEDEGAD
7	Áilen Niga Elle	GCGFFFG
8	Lemmon Elle	GGCCAAG

needed to find the key in Western music. Seven tones also gave a reasonable number of two- and three-tone transitions (six and five, respectively). The tone durations were not taken into account.

The first model correlated the distribution of the first seven tones with the tone distributions in the hymn and yoik corpora. The second model correlated the distribution of two-tone transitions in the initial segment with the two-tone transitions in the hymn and yoik corpora. Similarly, the third model correlated the distribution of three-tone transitions with the three-tone transitions in the two corpora. If the higher correlation is with the correct corpus statistics, we will say that the model has made a correct style classification. If not, it has made a classification error. Consequently, we will concentrate on the differences between the correlations with the two style statistics.

Figure 4 shows, for the three models, the difference between the correlation with the hymn statistics and the yoik statistics. As can be seen, the first model makes numerous classification errors. Correlations for correct classifications did not reliably exceed those for error classifications, $t(1,15) = .32$, *ns*. This result would be expected given the very similar tone distributions shown in Figure 3. The second model, using two-tone transitions,

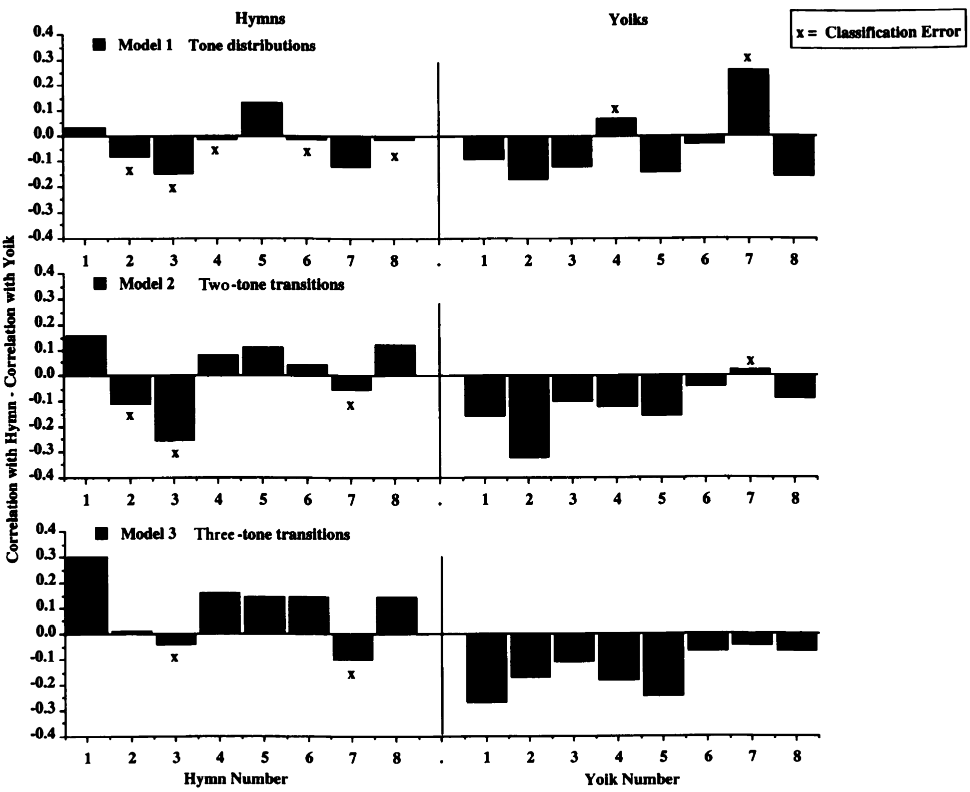

Fig. 4. Categorizations made by the three models. The values plotted are the correlation between the test material and the hymn corpus minus the correlation between the test material and the yoik corpus. Positive values mean that the model classified the test excerpt as a hymn, and negative values mean that the model classified the test excerpt as a yoik. Model 1 correlated the test excerpts with the tone distributions. Model 2 correlated the test excerpts with the two-tone transitions. Model 3 correlated the test excerpts with the three-tone transitions. Errors are indicated by x's.

made fewer classification errors. With some notable exceptions (Hymns 2, 3, and 7), the second model discriminates better between the styles, and the correlations for correct classifications exceeded those for error classifications, $t(1,15) = 2.1$, $p = .03$, one-tailed. Finally, the third model shows the best discrimination between the styles. Only Hymn 7 is strongly misclassified. Overall, the three-tone model performs quite well, with the correlations for the correct classifications higher on average than those for the incorrect classifications, $t(1,15) = 4.4$, $p = .0003$). In sum, the results of these analyses are that the algorithm's ability to discriminate style increases from the tone distribution model to the two-tone model to the three-tone model. We now consider the differences in the sequential order of tones that distinguish between the hymns and the yoiks.

Sequential Order in the Hymns and Yoiks

The following analysis was done to compare the two-tone and three-tone transitions in the hymns and yoiks. We consider the two-tone transitions first. The frequency of all two-tone transitions in the yoiks was subtracted from the frequency of the corresponding two-tone transitions in the hymns (after both sets of values were normalized). The 10 biggest differences on both ends of the scale were selected for presentation. Figure 5 shows on the left the 10 two-tone transitions that appeared in the hymns far more frequently than in the yoiks. The dashed lines represent two-tone transitions that are scalar movement. As can be seen, the hymns are characterized by scalar movement, with the exception of the BD sequence only. Figure 5 shows on the right the 10 two-tone transitions that appeared in the yoiks far more frequently than in the hymns. The circles here represent repeated tones, for example, CC. Here, tone repetition is far more frequent than in the hymns, and only two of the sequences, CD and DC, are scalar movement. The pentatonic orientation of the yoiks is apparent in that none of the frequent two-tone transitions includes F or B. Nor does A appear frequently in two-tone transitions. The interpretation of the yoiks as having two reference pitches is supported by the frequent succession of CG. These two reference pitches appear to be reinforced by movement to and from D and E, respectively.

Figure 6 shows the same analysis done for the three-tone transitions. Scalar movement in the hymns is even more clear here than for the two-

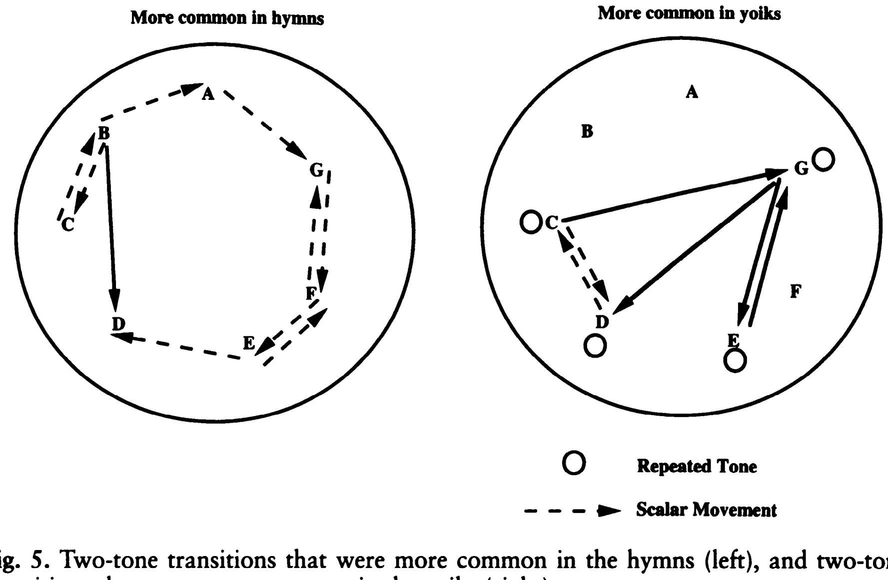

Fig. 5. Two-tone transitions that were more common in the hymns (left), and two-tone transitions that were more common in the yoiks (right).

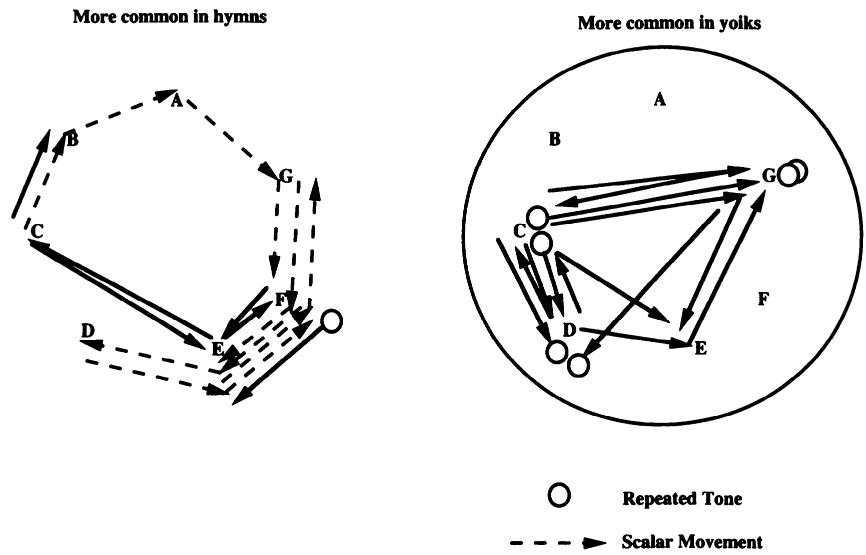

Fig. 6. Three-tone transitions that were more common in the hymns (left), and three-tone transitions that were more common in the yoiks (right).

tone transitions. Many of the patterns in the hymns consist of successive steps along the scale, for example, DEF, EFG, CBA, and so on. Two other cases involve alternations with E, CEC and FEF. Finally, two cases involve repetition followed by a scalar step, CCB and FFE. The yoik results, shown on the right, contain no movement of three tones along the scale. The relatively frequent sequences involving C and G are again apparent, as are the sequences involving C and D and the sequences involving G and E. Again, the prevalence of tone repetitions is seen, especially in the GGG sequence represented by the double circle.

With these results in mind, we can turn to the cases that were problematic for the algorithm. The first seven tones of Hymn 2 are CCEGGAB. That this was classified as a yoik can be understood by the following factors: the repeated C and the repeated G, which are relatively frequent in the yoiks, and the fact that the only scalar movement, GAB, is not one of the most common scalar sequences in the hymns. The second problematic hymn was Hymn 3, in which the first seven tones are GGECEGG. Note here the frequent repetition of the G and the absence of scalar movement, both of which are typical of the yoiks. Also the excerpt does not contain any of the tones F, A, or B, a feature that is also typical of the yoiks. The next problematic case was Hymn 7 with the first seven tones GCCCDED. Here the relevant factors seem to be the repeated C, the GC sequence, the absence of F, A, and B, and the fact that the only scalar movement is C D E (which is not one of the three-tone transitions that strongly distinguishes between hymns and yoiks). Finally, Yoik 7 was not strongly classified as either hymn or yoik. Its first seven tones are GCGFFFG. Yoik 7 has some features in common with the hymns: the appearance of the F, the repetition of the F, and the alternation between F and G. However, it also has important features in common with the yoiks, namely, the sequences CG and GC. Thus, the algorithm's difficulty with this yoik can be understood.

General Discussion

To summarize first the results of the algorithm described here, it was based not only on the relative frequencies of tones (as was done in the earlier Krumhansl-Schmuckler algorithm; Krumhansl, 1990), but also on the frequencies of two- and three-tone transitions in the initial segments of the music. The algorithm was tested by asking it to distinguish between two musical styles using only the first seven tones of eight Finnish spiritual folk hymns and eight North Sami yoiks. The relative frequencies of the tones are quite similar in these styles. As would thus be expected, the algorithm based only on the relative frequencies of tones was largely unable to classify the excerpts correctly. However, greater accuracy in distinguishing between the styles was found when two- and three-tone transitions were considered. Comparing the frequent two- and three-tone transitions in the styles yielded information about how the styles differ. The hymns are distinguished by a diatonic tonality and frequent scalar movement. The yoiks, in contrast, appear to have a pentatonic orientation (in which the A is relatively weak), and two reference pitches, C and G, that are frequently sounded in succession. Scalar movement is far less common than in the hymns. These characteristics could be used to understand which excerpts presented difficulties to the algorithm. They also are suggestive of further theoretical developments, especially of the yoik style, which has not received extensive theoretical analysis.

At a more general level, this article was motivated by examining how two basic psychological principles are used and elaborated in music. The first is the existence of cognitive reference points. These have been established by a variety of experimental results summarized earlier. Cognitive reference points also have been shown to have statistical support in the frequency of reference tones in music. The present analyses of two- and three-tone transitions extends these observations by considering patterns of tone movement to and from the reference pitches, and in the case of the yoiks, between the two apparently equally strong reference pitches. The second principle is sensitivity to frequency information and use of that sensitivity in learning and classification. Previous experimental results (Krumhansl et al., 1999; Krumhansl et al., 2000) showed that musical expertise was associated with melodic expectations based on sensitivity to longer sequences. In a complementary way, the algorithm that took into account sequential order performed more accurately in distinguishing between the two musical styles. In closing, let us consider two recent experimental results demonstrating the importance of sequential information in other communication systems.

A recent paper (Aslin, Saffran, & Newport, 1998) is concerned primarily with language acquisition but has intriguing parallels with music. Aslin et al. reported that 8-month-old infants can segment a continuous stream of speech syllables into wordlike units after only 3 min of exposure. This ability appears to derive from their sensitivity to the successive ordering of syllables. Aslin et al. created an artificial language of trisyllabic nonsense words; one language consisted of the words: pabiku, tibudo, golatu, daropi. The familiarization period consisted of a continuous stream of these trisyllabic nonsense words. The words were heard in random order with no breaks between them, such as "pabikugolatudaropitibudodaropigolatu...." After the familiarization period, which lasted 3 min, infants were presented with words and part-words (trisyllabic sequences spanning word boundaries). Their behavioral results showed that they discriminated between trisyllabic words and part-words. The words and part-words differed in the transition probabilities. The transition probabilities between the syllables were higher in the words than in the part-words. In other words, infants appear to be sensitive to the successive ordering of syllables in the artificial language. Thus, the research suggests that in both language and music, listeners may be sensitive to order information and use this sensitivity to form larger units in perception and memory.

An interesting application of sequential probabilities also appears in a recent study of bird songs (Gentner & Hulse, 1998). Bird songs serve at least three functions (which are different for different bird species and are not mutually exclusive). First, songs can be used by males to attract mates. Second, songs can be used to defend the nest site. Third, songs can be used to recognize individual birds, such as a female recognizing its mate. Their study focused on the last function, individual recognition, and examined what about an individual bird's song might carry the information about the identity of the singer. They used the following terminology to describe songs. A song consists of a number of sequentially patterned note clusters, called "song types." These are generally less than 1 s long and are often repeated before the next song type is sung. Sequences of song types are strung together in time to produce what is called a "song bout."

Apparently, different individual birds string together the song types in different ways. Gentner and Hulse (1998) did a statistical analysis of the frequencies of successive song types in European starlings. The first analysis describes the frequency distribution of song types, that is, how often different song types occur. The second analysis describes the frequency of successive pairs of song types (like two-tone transitions in music). The third analysis describes the frequency of successive triples of song types (like three-tone transitions in music). They then synthesized song bouts that conformed to these different statistics. That is, they took the statistics from different birds and created new song bouts in accordance with the relative frequencies of song types, two-song type sequences, and three-song type sequences.

Three groups of birds were first trained. The first group consisted of males who learned (through reinforcement) to identify their own songs from those of four familiar males. The second group was familiar with the songs, but none of the songs was their own; they were trained to identify the songs of a particular bird. The third group consisted of females who were previously unfamiliar with all the songs; they were also trained to identify the songs of a particular bird. The question, then, was whether the trained birds could generalize this training to the novel synthesized song bouts, correctly identifying the individual. The training successfully transferred to the novel song bouts. Moreover, the transfer was more successful for the synthesized song bouts that matched the relative frequencies of longer sequences of song types. These results suggest that birds, as well as humans, are sensitive to the successive ordering of sound events. In bird song, it appears to be used to identify the individual birds, at least by European starlings.

Thus, sensitivity to the order of events appears to play a role in a variety of communicative contexts. It appears to be a basic psychological principle in both humans and other species. As a consequence of this generality, we would expect to find its influence in music cross-culturally. This has been shown not only in the behavioral experiments in previous research, but also in the statistical style analysis summarized here. The algorithm showed the utility of sequential information in classifying musical styles. This general principle of sensitivity to sequential order is elaborated in each of these systems in different ways. In music, it is coordinated with reference pitches, scale structure, and harmony, and it is elaborated by rhythm and phrasing. In language, it appears to yield information about the sequences of sounds that form words of the language and possibly also syntactic categories. In animal communication, there is at least one example of sequential ordering of sound types yielding identification of individuals. Despite these differences, sensitivity to sequential ordering appears to be a core principle in encoding and interpreting sequences of sounds.[1]

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*, 321–324.

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

1. The research in Finland was supported by a Fulbright Fellowship awarded to Carol Krumhansl and a grant from the Finnish Academy awarded to the Cognitive Musicology group at the University of Jyväskylä.

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

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

Gentner, T. Q., & Hulse, S. W. (1998). Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris. Animal Behavior, 56,* 579–594.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Kohonen, T. (1997). *Self-organizing maps* (2nd ed.). Berlin: Spring-Verlag.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11,* 346–374.

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

Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum, 17,* 53–80.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89,* 334–368.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579–594.

Krumhansl, C. L., Louhivuori, J., Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectancy in Finnish folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception, 17,* 151–195.

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T. & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition, 75,* 1–46.

Oram, N., & Cuddy, L. L. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research, 57,* 103–118.

Rosch, E. (1979). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press.

Rowe, R. (1993). *Interactive music systems: Machine listening and composing.* Cambridge, MA: MIT Press.

Rowe, R. (2000). Key induction in the context of interactive performance. *Music Perception, 17,* 511–530.

Schmuckler, M. A. (1989). Expectation in music: Investigations of melodic and harmonic processes. *Music Perception, 7,* 109–150.

Vos, P. G. (2000). Tonality induction: Theoretical problems and dilemmas. *Music Perception, 17,* 403–416.

Vos, P. G., & Van Geenen, E. W. (1996). A parallel-processing key-finding music. *Music Perception, 14,* 185–223.