Melodic Contour Similarity Using Folk Melodies

Mark. A. Schmuckler
University of Toronto Scarborough, Scarborough, Canada

melodic contour, or the pattern of rises and falls in pitch, is a critical component of melodic structure, and has an important impact on listeners' perceptions of, and memory for, music. Despite its centrality, few formal models of contour structure exist. One recent exception involves characterizing contour by the relative degrees of strength of its cyclic information, quantified via a Fourier analysis of the pitch code of the contour. Three experiments explored the applicability of this approach, demonstrating that listeners' similarity ratings for pairs of melodies were predictable from Fourier analysis quantifications of rhythmically complex (Experiment 1) and rhythmically simple (Experiment 2) melodies, as well as for derived similarity measures based on melodic complexity judgments (Experiment 3). These findings indicate that Fourier analysis is an effective model of melodic contour, and that it can predict perceived melodic similarity.

Received January 29, 2009, accepted April 1, 2010.

Key words: melodic contour, melodic perception, melodic similarity, Fourier analysis, time series analysis

Over the years, research in music cognition has elucidated a handful of factors that play a critical role in listeners' apprehension of complex auditory information. Within the pitch domain, two factors consistently provide a focal structure for music perception (Dowling, 1978; Schmuckler, 2004, 2009)—the perception of tonality, or musical key (see Justus & Bharucha, 2002; Krumhansl, 1990, 2000a; Schmuckler, 2004; Schmuckler & Tomovski, 2005, for reviews), and the apprehension of melodic contour (see Justus & Bharucha, 2002; Quinn, 1999; Schmuckler, 1999, 2004; 2009, for reviews). Tonality refers to the structuring of a set of pitches around a central reference pitch, such that these tones tend to be heard in relation to this pitch (Krumhansl, 1990, 2000a; Schmuckler, 2004; Schmuckler & Tomovski, 2005; Smith & Schmuckler, 2004). Contour refers to the pattern of rises and falls along an auditory or musical dimension (Dowling, 1971, 1978, 1994, 2001; Dowling & Fujitani, 1971; Dowling & Harwood, 1986; McDermott, Lehr, & Oxenham, 2008; Schmuckler, 1999, 2004; Schmuckler & Gilden, 1993). Most typically, contour is applied to changes in pitch information, although it is relevant to other auditory dimensions such as loudness, timbre, and rhythm (Marvin, 1991, 1995; McDermott et al., 2008; Schmuckler & Gilden, 1993).

Given their respective roles in musical processing, a number of researchers have proposed models for the apprehension and cognitive structure of both parameters. By and large, the majority of attention has been devoted to models of tonality and the process of key-finding (Brown, 1988; Brown & Butler, 1981; Brown, Butler, & Jones, 1994; Browne, 1981; Butler, 1989; Krumhansl, 1990, 2000b; Krumhansl & Schmuckler, 1986; Krumhansl & Toiviainen, 2001; Leman, 1995; Schmuckler & Tomovski, 2000, 2005; Smith & Schmuckler, 2004; Temperley, 1999, 2001, 2007; Tillman, Bharucha, & Bigand, 2000; Toiviainen & Krumhansl, 2003).

In contrast, far fewer models have been proposed for understanding the structure and perception of melodic contour (Adams, 1976; Eerola, Himberg, Toiviainen, & Louhivuori, 2006; Friedmann, 1985, 1987; Marvin, 1991, 1995; Marvin & Laprade, 1987; Morris, 1993; Quinn, 1999; Schmuckler, 1999). On a music-theoretical basis, probably the most common form of contour model has been what is called a combinatorial model (Polansky, 1987; Quinn, 1999). In such a model, each note of the melody is judged relative to its pitch relations with all other notes in the melody. This entire set of two note pitch relations can then be characterized by coding ascending pitch relations with a value of 1, equivalent pitch relations with a value of 0, and descending pitch relations with a value of –1 (although other coding systems are possible; see Quinn, 1999, for a variant). These models typically result in a matrix of pitch height relations between all possible combinations of adjacent and nonadjacent notes fully representing the interval content of a melody. Such matrices can then be used to predict the theoretical similarity between melodies, based on the overlap of the content of these representations.

Music Perception volume 28, issue 2, pp. 169–193, issn 0730-7829, electronic issn 1533-8312 DOI:10.1525/mp.2010.28.2.169

Such models have only proven moderately successful in terms of characterizing actual perceived contour similarity, however. For instance, Schmuckler (1999)—which will be described in greater detail subsequently—found no evidence for any of the myriad of possible measures produced by such combinatorial models in predicting listeners' perceived contour similarity. In contrast, Quinn (1999) did find that for a set of short, specially composed melodies, quantifying the overlap in interval content between melody pairs predicted listeners' subsequent similarity ratings of these melodic pairs. Of particular importance for these predictions was the interval overlap between contiguous, or successive notes, and to a lesser extent, the interval overlap between non-contiguous notes separated by a single note (e.g., the interval between the first and third notes, or the second and fourth). This prediction, based on contiguous note information, is of particularly relevance given that this information most closely parallels the pattern of rises and falls in pitch of the actual musical surface, or the melody itself.

One limitation to combinatorial models of pitch contour is that they are fundamentally local models based on aggregated pitch relations between individual pairs of notes. Unfortunately, such models fail to capture the global pattern of rises and falls in a melody, an aspect that intuitively seems important for listeners' percepts of contour. Accordingly, Schmuckler (1999) suggested that this global sense of contour could be captured via a time series analysis of a melody. Specifically, Schmuckler proposed that a melody be coded in a simple 0 – N format (with N equal to the number of discrete frequencies occurring in the melody), and that this integer code could then be Fourier analyzed to determine the relative strengths of the cyclic components of the melody's contour. Fourier analysis is a mathematical procedure by which a complex signal, typically in the temporal or spatial domain, is converted into the frequency domain by decomposing the waveform into a set of harmonically related sine waves representing the presence of cyclical pattern information across the length of the original series. These sine waves are then characterized by their relative amplitude (strength) and phase (timing) relations, with the harmonic number of the sine corresponding to the frequency or the number of repetitions of that wave that occurs throughout the course of the series being analyzed. Thus, the first harmonic produced by the Fourier analysis represents a sine wave with only a single cycle throughout its length; the second harmonic represents two complete sine waves throughout the length of the series, and so on. As such, all of the harmonics describe the entire series being analyzed, while differing systematically (and as function of harmonic number) in the number of repetitions of up-down cycles (characteristic of a sine wave in zero phase) present in the waveform. When applied to a melody, Fourier analysis provides a convenient description of the contour – one that simultaneously takes into account both rapid, high frequency point-to-point fluctuations, as well as slower, low frequency trends.

Schmuckler (1999) went on to suggest that this contour description could be used to predict contour relations, such as perceived melodic complexity or contour similarity. In a test of this idea, Schmuckler gathered ratings of contour complexity and used these ratings to generate derived contour similarity scores. These derived similarities were then predicted by the degree of overlap of the amplitude and phase spectra of the contour pairs. Across a pair of experiments, Schmuckler found that derived contour similarities of a set of 12-tone rows and 12 note tonal melodies were predictable from the overlap, assessed by both correlation and difference score measures, of the melodies' amplitude spectra. Phase spectra overlap was more variable in its predictive power, correlating with derived similarity for the simple tonal melodies but not for the atonal 12-tone rows. Taken together, these findings thus demonstrated that listeners are sensitive to the amplitude (i.e., strength) and possibly the phase (i.e., timing) information of the cyclic components of melodies, and use this information in their judgments of contour similarity. More generally, these findings provide evidence that Fourier analysis can characterize melodic contour, both structurally and psychologically.

Schmuckler (1999) does not provide the only evidence attesting to a role for time series information in characterizing melodic contour. Eerola and colleages (Eerola & Bregman, 2007; Eerola et al., 2006), for instance, have examined the role of time series measures in predicting listeners' percepts of melodic complexity (Eerola et al., 2006) and melodic similarity. Eerola et al. (2006), for instance, found that the Fourier transform of a melody, as well as its autocorrelation function (which assesses the degree of self-similarity within a melody) predicted melodic complexity ratings for sets of African and Western folk melodies. In a subsequent follow-up study, Eerola and Bregman (2007) compared listeners' similarity judgments for pairs of 10-note melodies to a variety of melodic predictors, including one measure that encapsulated contour periodicity (as assessed via Fourier analysis). These authors again found that contour periodicity, among other factors, predicted similarity ratings, although these findings are difficult to interpret given the observed high degree of collinearity amongst the factors themselves. Together, these studies demonstrate the potential for Fourier analysis in characterizing melodic contour, and also for assessing contour similarity.

Recently, Prince, Schmuckler, and Thompson (2009a) demonstrated that observers' percepts of crossmodal (auditory and visual) contours also can be predicted on the basis of time series information. In this work, listeners rated the similarity of both short and long melodic contours presented as melodies and as line drawings. These ratings were predictable from similarity measures derived from the overlap of the surface structure, and the strength (amplitude) and timing (phase) of cyclical information as assessed by a Fourier analysis. Accordingly, these findings affirm the applicability of this type of contour model to crossmodal contexts.

Despite being compelling in many respects, these studies have some important limitations. One of the most significant concerns is the somewhat artificial nature of the melodic stimuli employed in these investigations. Eerola and Bregman (2007), for instance, used extremely short melodic patterns (from 4 to 9 notes), and it is not clear how well these fragments truly represented coherent musical phrases. Similarly, Schmuckler (1999), although choosing stimuli meant to be representative of 20th century and simple tonal music, nevertheless employed simplistic equitemporal contours. Such a choice simplified the application of the Fourier analysis model, but unfortunately limited the generalizability of this model. Of the studies that have employed more naturalistic contours, there also remain some concerns. Somewhat obviously, Prince et al.'s (2009a) primary goal was investigating crossmodal contour similarity, and as such did not test melodic contour similarity per se. Eerola et al. (2006), although also employing realistic melodies, did not actually test perceived contour similarity. Instead, these authors examined melodic complexity, a related but clearly distinct concept. Thus, although such work does imply that the Fourier analysis model can capture important melodic information, it is still an open question as to whether this approach provides a reasonable description of perceived similarity for more musically realistic contours.

A final concern with this previous work, and one exclusively applicable to Schmuckler (1999), arises from the use of a derived similarity value as the principal dependent measure. In Schmuckler (1999) this choice was primarily pragmatic, based on pilot results that were discouraging regarding the viability of a more direct similarity measure. Unfortunately, this decision does leave open the question of whether the Fourier analysis approach is effective in predicting more direct similarity judgments. Given all of these concerns, the current project examines the effectiveness of the Fourier analysis model in predicting listeners' perceptions of rhythmically complex melodic pairs, using a direct rating of perceived similarity.

Along with extending the assessment of the Fourier analysis model, a second goal of this project involves investigating alternative characterizations of melodic contour. As has been recognized by numerous authors, melodies and melodic contours can be described in a multitude of ways, thus potentially containing multiple components. Accordingly, any or all such components could reasonably play a role in perceived contour similarity, with the relation between these factors ranging from truly competitive, with one or the other model "winning" in its predictions of contour similarity, to complementary, with multiple factors mutually contributing to a more thorough characterization of contour similarity. Based on previous work, it seems most likely that it is the complementary, and not the competitive, scenario that best describes contour perception. Both Eerola and colleagues (Eerola & Bregman, 2007; Eerola et al., 2006; Eerola, Järvinen, Louhivuori, & Toiviainen, 2001) and Schmuckler and colleagues (Prince et al., 2009a; Schmuckler, 1999) have in their work found evidence for multiple factors in predicting contour similarity. Accordingly, it would be surprising if, within the current context, there was not similar evidence for an array of components in contour processing.

What are some of the potential components that might be relevant in this project? As already discussed, one class of models derives from time series, and specifically Fourier analysis, models of contour structure. Details on different forms of characterizations within this class of models will be provided subsequently.

Along with Fourier analysis models, previous work suggests a potential role for at least two other classes of models. The first is also based on Schmuckler's (1999) findings, in which the similarity between melody pairs was predictable based on an oscillation model of contour. In general terms, the oscillation model captures the degree of up and down movement contained in a contour. Schmuckler found that various means of characterizing such oscillation information were reliable predictors of contour similarity. Again, details on this information will be developed subsequently.

A final category of model involves similarity based on the degree of comparability in the surface characteristics of the various melodies. Such surface information assesses the relative positioning of peaks and troughs in the contours themselves, and has been found to play a role in perceived similarity (Prince et al., 2009a; Quinn, 1999). One such source of evidence for this factor stems from Quinn's quantifications of the music-theoretic models of Marvin and Laprade (1987). According to Shmulevich (2004) a simple correlation of melodic surface information is actually equivalent to a core component of Quinn's music-theoretic model. In addition, in their cross-modal investigations, Prince et al. (2009a) found that the surface similarity between visual and auditory contours predicted observers' perceived similarity ratings. Accordingly, there is ample evidence to suggest a role for surface information in contour similarity. As with the previous factors, more detailed discussion of different forms of surface information is provided subsequently.

Experiment 1: Similarity of Rhythmically Diverse Folk Melodies

The focus of Experiment 1 was to test the various classes of models just described, employing a set of more realistic musical stimuli, and using a direct similarity rating procedure. Accordingly, this experiment employed a set of folk melodies varying in terms of the number of notes contained in each melody, the rhythmic pattern of these notes, and so on; all parameters that were controlled (and equated) in Schmuckler (1999). Moreover, in this experiment listeners heard all possible pairings of these twelve melodies, and provided direct contour similarity ratings. As such, this study represents a significant extension of these earlier findings, and a strong test of the predictive validity of all of the models.

Method

PARTICIPANTS

The sample of listeners consisted of sixteen undergraduate students (*M* age = 19.6 years, *SD* = 1.5), who received either course credit in introductory psychology or $7 for participating. One additional listener was run, but her data were lost due to a computer error. All listeners were musically trained, with a mean of 11.0 (*SD* = 3.8) years playing an instrument or singing, a mean of 3.2 (*SD* = 3.4) hours/week currently involved in music making activities, and a mean of 14.9 (*SD* = 11.2) hours/week listening to music. All listeners reported normal hearing.

EXPERIMENTAL APPARATUS AND STIMULI

Stimuli were generated using a Yamaha TX816 synthesizer, connected to an IBM-compatible 286 MHz computer by a Roland MPU-401 MIDI controller. The timbre employed in this experiment was harmonically complex, and approximated the sound of a piano (Schmuckler, 1989). All tones were input into a Mackie 1202 mixer, then amplified and presented to listeners over a Boss MA-12 micromonitor speaker, at a comfortable listening level.

The stimuli for this study consisted of 12 folk melodies, shown in Figure 1. Each melody was four measures in length, not including an initial "pick-up" beat, with the duration of each beat set to 500 ms. All melodies were played in the same general pitch range, and instantiated either a G major or E minor tonality; these two keys are strongly musically related.

PROCEDURE

Listeners were told they were participating in an experiment on contour perception. They were informed that on each trial they would hear a pair of melodies, and that they were to rate the similarity of the melodies based on their melodic contour. Melodic contour was described as the pattern of rises and falls in pitch over the course of the melody. These ratings were to be made on a 9 point scale, with 1 indicating very different contours and 9 indicating very similar contours.

On each trial listeners heard one melody, followed by a 500 ms pause, and then the second melody. After the second melody, listeners typed in their rating on the computer keyboard and the computer automatically began the next trial. At the beginning of the experiment listeners received a practice block of 12 trials, with these trials consisting of all possible combinations of four simple melodic patterns based on a G major scale. After completing these practice trials, listeners immediately began a block of 132 experimental trials. This block consisted of all possible counterbalanced combinations of the 12 melodies of Figure 1 (e.g., melody 1 followed by melody 2; melody 1 followed by melody 3 . . . melody 3 followed by melody 1; melody 2 followed by melody 1), excluding repetitions of the same melody. Subsequently, listeners completed a musical background questionnaire and were debriefed as to the purposes of the experiment. The entire experimental session lasted 30 – 45 minutes.

Results

As a prelude to analyzing listeners' perceived similarity ratings, the various theoretical models of contour similarity described earlier were quantified. These models were then compared to perceived similarity values, to assess the efficacy of these earlier models.

THEORETICAL MODELS OF CONTOUR SIMILARITY

Fourier analysis model. The first predictor of contour similarity was based on the Fourier analysis model. Investigating rhythmically complex, unequal length contours such as those used in this study provides interesting challenges for the application of the Fourier analysis model. One such challenge involves the format of the numeric code used to represent the individual pitches of the melodies. In Schmuckler (1999), contours were coded in 0 to N – 1 format, with N equal to the highest distinct frequency.

FIGURE 1. The folk melodies employed as experimental stimuli in Experiments 1 and in the rhythmic condition of Experiment 3.

Because melodies were of the same length and were equitemporal, there was no concern over how to incorporate rhythmic variation into these contour codes. Unfortunately, by employing melodies in which the component notes are no longer equitemporal, the issue of rhythmic variation cannot be so easily circumvented. As such, the question of how best to account for such variation in an assessment of contour similarity becomes critical.

In considering this issue, there are two possibilities. On the one hand, it could be argued that the rhythmic information contained in a melody should be treated as an independent factor in assessment of contour similarity. Such an approach is justified based on evidence for the independent processing of pitch and rhythm information in melody perception (e.g., Monahan, 1993; Palmer & Krumhansl, 1987a, 1987b; Thompson, Hall, & Pressing, 2001; Thompson & Sinclair, 1993). On the other hand, it is possible to include rhythmic variation directly into the pitch code itself. This approach is conceptually aligned with the idea that pitch and temporal information are interactive in various aspects of melodic processing (Boltz, 1991, 1993; Boltz & Jones, 1986; Jones, 1993; Jones & Boltz, 1989; Jones, Boltz, & Kidd, 1982; Jones & Pfordresher, 1997).

Unweighted pitch code	1	2				0	
Rhythmic code	2	4				1	
Weighted pitch code	1	1	2	2	2	2	0

FIGURE 2. A sample three tone melodic contour, along with various forms of pitch codes (unweighted and weighted) and the rhythm code.

To make these considerations concrete, take the three-note melodic fragment shown in Figure 2. In terms of its relative pitch content, this fragment contains a note of medium frequency, followed by a note of higher frequency, followed by a note of lower frequency. In terms of its duration content, the second note is twice as long as the first note, and the third note is half as long as the first note. If one wished to code the pitch content only for this fragment, its integer representation would be 1 2 0, with 1 representing the first note, 2 the second note, and 0 the third note. This representation, which appears in Figure 2 as the *unweighted pitch code*, codes relative pitch variation only, not absolute frequency differences between notes. This code has the advantage that it portrays frequency change only, and gives a "purer" representation of the rises and falls in pitch contour. In fact, this is the type of code that has been typically adopted in contour analyses (e.g., Friedmann, 1985, 1987; Marvin, 1991, 1995; Marvin & Laprade, 1987; Morris, 1993), as well as in previous studies on perceived contour similarity (Quinn, 1999; Schmuckler, 1999). Figure 2 also shows a rhythm code for this fragment, labeled *rhythm code*, which consists of an integer representation of the relative durations of these tones, coded in terms of the shortest duration unit, which is arbitrarily given a value of 1. Finally, Figure 2 presents a code that incorporates both pitch and rhythm into a single contour representation, with each pitch element weighted by its duration; this code is labeled as the *weighted pitch code*. One advantage to this coding scheme is that it provides a literal representation of what is heard by listeners in terms of how pitches change in real time. As an aside, it is important to note that the terms "weighted" and "unweighted" are not being used in the sense often employed in psychological and specifically connectionist contexts, in which certain values (in this case the pitch codes) are literally multiplied by the weighting factor (the rhythm code). Instead, in the current context, weighting is being used to notate a contour code that is either elaborated by each pitch's durational component (the weighted code) or is not durationally elaborated (the unweighted code). For simplicity in presentation, this concept of durational elaboration will be referred to as weighted versus unweighted.

The question of whether an unweighted or weighted pitch code is most appropriate is further complicated by issues related to how contours of different lengths can be compared. For unweighted pitch codes, if two contours vary in the number of their constituent notes, their pitch codes will correspondingly vary, and most importantly, their respective amplitude and phase spectra also will vary in the number of harmonics produced. Accordingly, comparison of sets of spectra will be constrained by the number of harmonics of the shorter of the two melodies,[1] comparing the strength and timing of harmonics representing the same frequency of repetitions within the melody. Even more importantly, comparisons of such spectra now reflect similar cyclic patterns in terms of the melody as the unit of analysis. For instance, based on this analysis, one could say that melody x is characterized predominantly by three cyclic repetitions within its length, whereas melody y is characterized primarily by four cyclic repetitions over the course of its length.

In contrast, if the contours are of equal length in terms of the number of musical beats each encompasses (e.g., both melodies are, say, four measures long), a weighted pitch code equalizes the number of contour elements, even if the melodies differ in their actual number of notes. Correspondingly, it is thus straightforward to compare the resulting amplitude and phase spectra produced by a Fourier analysis. Moreover, contour characterizations now reflect cyclic pitch structure in time, relative to note durations. Thus, comparisons of the relative efficacy of unweighted versus weighted pitch codes in predicting perceived similarity provide a means for determining the relative importance of duration information in contour characterizations.

Given these considerations, two different pitch codes were generated for each melody. In the *unweighted pitch code* each note was represented using an integer between 0 and N – 1, with N representing the number of unique pitches in the melody. In the *weighted pitch code*, the 0 to

[1]An alternative solution to this problem would be to equate the number of elements of the two pitch codes by padding the shorter code with the mean value of the sequence; comparison of Fourier spectra would now involve sequences of equal length, and no longer require removing upper frequency information. In fact, this procedure produces similar results to that provided by removal of the upper frequency information, as described in the text. Accordingly, the conceptually simpler method of removing the upper frequency information was employed in this work.

M01

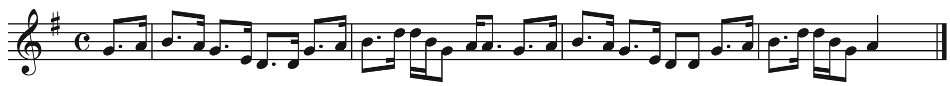
Pitch Code:

Unweighted: 1 2 3 3 2 1 2 2 3 2 1 2 3 3 2 2 1 2 3 2 1 0 1

Weighted: 1122 3333333 22221111 22222222 33221122 33333333 22221122 33221100 1111

Rhythm Code: 2 2 6 2 4 4 4 4 2 2 2 2 6 2 2 2 2 2 2 2 2 2 4

M05



Pitch Code:

Unweighted: 2 3 4 3 2 1 0 0 2 3 4 5542 33 2 3 4 3 2 1 0 0 2 3 4 5542 3

Weighted: 2223 44432221 00002223 44455422 33332223 44432221 00002223 444554223

Rhythm Code: 3 1 3 13 1 3 13 1 3 1112 13 3 1 3 13 1 2 2 3 1 3 11124

M07

Pitch Code:

Unweighted: 0 1 2 22 0 3 33 2 2 22 0 1 0 1 2 22 0 3 45 5 4 3221 0

Weighted: 0001 22222220 333333322 22222220 11110001 22222220 33345555 44432211 0000

Rhythm Code: 3 1 3 13 1 3 13 1 3 13 1 4 3 1 3 13 1 3 13 1 3 1112 4

Surface Correlations

Contours	Unweighted Pitch Code	Weighted Pitch Code	Rhythm Code
M01 – M05	0.33	0.11	-0.02
M01 – M07	0.07	0.12	0.14
M05 – M07	-0.49	-0.07	0.52

FIGURE 3. The unweighted pitch code, weighted pitch code, and rhythm code, as well as the surface correlations for these codes, for three sample stimuli of Experiment 1.

N – 1 code for each note was elaborated by its duration, using a one-quarter subdivision of the beat as the unit of analysis. Figure 3 presents three of the stimulus contours from this study, along with their unweighted and weighted pitch codes. Figure 3 also shows the duration values for each note used to create the weighted pitch code; these numbers, which will be discussed shortly, made up the *rhythm code* of these melodies.

Amplitude and phase spectra were calculated by Fourier analyzing the unweighted and weighted pitch codes; Figure 4 presents these spectra for the analysis of the unweighted pitch codes of the sample melodies of Figure 3. The degree of contour similarity between all pairs of melodies (based on amplitude and phase spectra) was then calculated by aligning the amplitude and phase spectra for each pair of contours and computing the mean absolute difference between these spectra values.[2]

[2]Along with difference scores it is possible to correlate the amplitude and phase spectra. In fact, previous work employing both measures (Prince et al., 2009a; Schmuckler, 1999) has found that both indexed perceived similarity ratings, although difference score values appear to be more sensitive than do correlation measures. With regard to the current data, both measures were calculated, and by and large produced comparable results. Accordingly, only the difference score analyses will be described.

Fourier Components

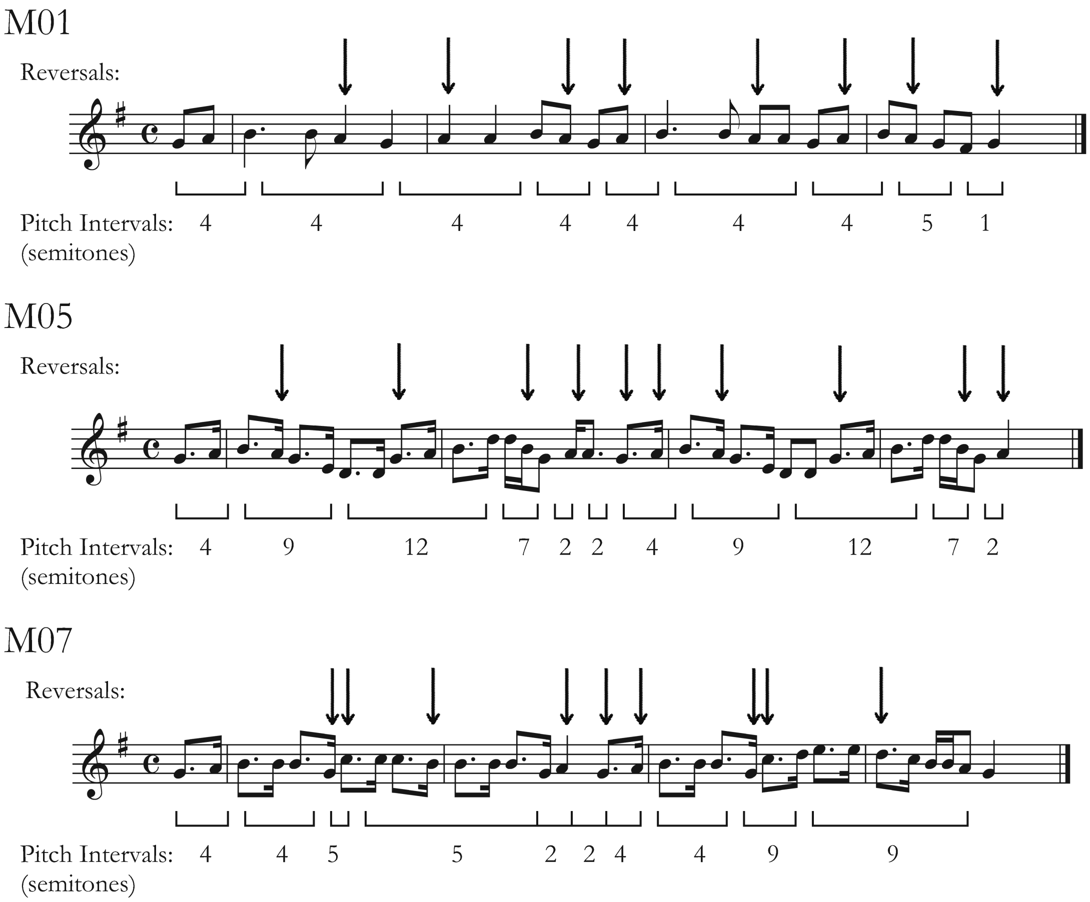
Harmonic	M01 Amplitude	M01 Phase	M05 Amplitude	M05 Phase	M07 Amplitude	M07 Phase
1	4.59	-2.65	0.74	-2.20	9.54	1.46
2	4.63	-1.77	21.92	1.15	21.76	3.13
3	4.75	-2.11	3.63	1.10	5.95	-2.20
4	8.71	-1.69	23.86	-2.09	8.58	-1.01
5	4.04	1.61	3.41	-2.01	3.26	1.80
6	1.99	0.96	1.51	2.73	10.56	-2.54
7	2.84	2.39	2.48	-2.10	2.44	2.83
8	1.59	-0.36	2.96	0.78	4.05	1.90
9	0.16	1.64	2.66	-1.22	3.44	-1.02
10	3.17	2.49	6.58	2.04	2.16	-0.03
11	0.25	1.59	1.00	3.14	1.45	-1.50
12			3.30	1.64	4.58	-2.33
13			0.89	1.02	4.21	1.17
14			3.18	2.56	4.66	2.64
15			2.14	2.33	4.13	0.43
16			1.11	-2.61		

Absolute Difference Scores

Harmonic	M01– M05 Amplitude	M01– M05 Phase	M01– M07 Amplitude	M01– M07 Phase	M05 – M07 Amplitude	M05 – M07 Phase
1	3.85	0.46	4.95	4.12	8.80	3.66
2	17.29	2.92	17.13	4.90	0.16	1.98
3	1.12	3.20	1.20	0.09	2.32	3.30
4	15.15	0.40	0.13	0.68	15.27	1.08
5	0.63	3.61	0.78	0.19	0.15	3.81
6	0.47	1.77	8.57	3.51	9.04	5.27
7	0.36	4.49	0.40	0.45	0.04	4.93
8	1.36	1.13	2.45	2.26	1.09	1.12
9	2.50	2.86	3.28	2.66	0.78	0.20
10	3.41	0.44	1.01	2.51	4.42	2.07
11	0.75	1.56	1.20	3.09	0.45	4.65
12					1.28	3.97
13					3.32	0.14
14					1.48	0.08
15					1.98	1.90
Mean Absolute Difference Score	**4.26**	**2.08**	**3.74**	**2.22**	**2.19**	**2.56**

FIGURE 4. Amplitude and phase spectra, along with amplitude and phase spectra difference scores, for three sample stimuli of Experiment 1.

For the unweighted pitch codes, difference scores were based on only those harmonics contained in both contours. Thus, comparison of contours M01 and M05 in Figure 4 was based on the difference in 11 harmonics, whereas comparison of contours M05 and M07 was based on 15 harmonics. Figure 4 also shows these difference scores for both amplitude and phase spectra for the three sample melodies. Because these melodies were all of equal total length in terms of the number of beats, Fourier analysis of the weighted pitch codes produced amplitude and phase spectra for each contour containing comparable numbers of harmonics. Accordingly, difference scores were calculated based on the full complement of amplitude and phase spectra information. Ultimately, these comparisons produced a set of half-matrices representing theoretical similarity between unweighted and weighted pitch codes, based on both amplitude and phase spectra differences.



Contours	Summed Reversals	Mean Reversals	Summed Interval	Mean Interval
M01	8	.35	34	3.78
M05	10	.31	70	6.36
M07	9	.29	48	4.80
Difference Scores				
M01 – M05	2	.04	36	2.74
M01 – M07	1	.06	14	1.00
M05 – M07	1	.02	22	1.56

FIGURE 5. Contour reversals (notated by arrows above the melodies) and pitch intervals (notated by brackets below the melodies) for three sample stimuli of Experiment 1.

Oscillation models. Four different oscillation measures were developed. The first two measures involved contour reversals, including a count of the total number of contour reversals occurring in the melody (i.e., changes in interval direction, such as ascending → descending or descending → ascending), and the mean number of contour reversals (the summed contour reversals divided by the number of notes in the contour). The final two measures involved the interval sizes within the contour, and included the summed interval size, which was the summed number of semitones contained within each individual contour segment (a single direction motion, either up or down contour segments), and the mean interval size, which was the summed interval size divided by the number of contour segments. Based on these oscillation measures, contour similarity can be calculated by creating difference scores between the various values. Figure 5 shows the reversal and interval size measures for the sample contours of Figure 3, and their respective similarities.

Surface correlation models. The final category of models involved similarity based on surface characteristics of the melodies. To assess surface similarity, the unweighted pitch codes, the weighted pitch codes, and the rhythm codes for the 12 melodic contours were intercorrelated. As with the previous analyses, when contours of differing length were compared, the number of elements entering into the correlation was based on the length of the shorter of the two contours. Figure 3 presents the surface correlations of the unweighted pitch code, the weighted pitch code, and the rhythm code for the three sample contours.

PREDICTION OF LISTENERS' CONTOUR SIMILARITY

Initial analyses explored the impact of melodic ordering by correlating listeners' similarity values for the two half-matrices representing the varying melodic ordering (e.g., melody M01 followed by melody M02 versus melody M02 followed by M01). On an individual listener basis, the degree of comparability between these two half-matrices was somewhat variable, with individual listener correlations ranging from –.07 to .54 ($M = .16$, $SD = .18$). Although potentially worrisome, it should be remembered that listeners received only a single exposure to each melodic pair in each ordering; accordingly, a fair degree of variability is to be expected. Of more interest is the correlation between the half-matrices corresponding to different melodic orderings produced by averaging listeners' similarity ratings; such values reflect similarity between melodic ordering pairs that are less influenced by individual variability. More reassuringly, these two half-matrices were strongly related, $r(64) = .55$, $p < .001$, indicating that similarity ratings were reasonably comparable regardless of melodic ordering.

Given the reasonable correspondence between half-matrices on an averaged listener basis, ratings of contour similarity were ultimately combined across these half-matrices for each listener, and then these averaged listener half-matrices were themselves averaged to produce a single half matrix reflecting the average perceived ratings of contour similarity between all possible pairs of melodies. This contour similarity half-matrix could then be predicted from the theoretical models just articulated, to determine which of these factors were related to contour similarity judgments.

Table 1 shows the simple correlations between the theoretical models and the listeners' similarity ratings, and presents a number of intriguing results. First, replicating previous findings (Eerola & Bregman, 2007; Eerola et al., 2006; Schmuckler, 1999), perceived contour similarity was predicted by differences in the amplitude spectra between melodies produced by the Fourier analysis model. It is interesting to note that this predictive relation

TABLE 1. Correlations Between the Theoretical Models of Perceived Contour Similarity and Listeners' Ratings of Contour Similarity for Experiments 1–3

Model	Experiment 1 Rhythmic Melodies	Experiment 2 Equitemporal Melodies	Experiment 3 Complexity Ratings		
			Rhythmic Group		
Pitch Code			Rhy	Eqi	Avg
Unweighted					
Amplitude	–.04	–.04	–.20	–.14	–.19
Phase	–.06	–.13	–.02	–.02	–.02
Weighted					
Amplitude	–.22 C	–.25*	–.40**	–.43**	–.47**
Phase	–.25*	–.26*	–.22C	–.24A	–.25*
Oscillation					
Summed Reversals	–.28*	–.08	–.10	–.05	–.08
Mean Reversals	–.22 C	–.29*	–.09	–.12	–.12
Summed Interval Size	–.02	–.04	–.04	.02	–.01
Mean Interval Size	–.08	–.06	–.36**	–.55**	–.51**
Surface Correlation					
Unweighted Pitch	.18	.07	.21	.17	.21
Weighted Pitch	.03	–.04	.03	.10	.08
Rhythm	.51**	.23 A	.29*	.23B	.29*

*$p < .05$ A $p = .06$
**$p < .01$ B $p = .07$
C $p = .08$

held for the weighted, but not the unweighted pitch codes. Similarly, the predictive power of phase spectra differences was also significant for the weighted but not the unweighted pitch codes.

In terms of the other models, there was some support for the oscillation model in the current results. Specifically, similarity values based on the summed number of reversals in the contour also significantly predicted perceived similarity, and the mean number of contour reversals was marginally related. In contrast, neither oscillation measures based on the interval sizes of the contour predicted listeners' ratings. These results provide a partial replication of Schmuckler (1999), who found that reversals did predict derived similarity measures (in one of two experiments), although they diverge from this earlier work in that Schmuckler also found that the summed interval size was a predictor of contour similarity. As for the surface correlation models, the simple correlations between the unweighted or weighted pitch codes both failed to predict listeners' perceived similarity. In contrast, however, the correlation between the rhythmic codes for these contours was a strong predictor of contour similarity.

Overall, there was positive evidence for a number of factors in predicting listeners' perceived contour similarity, including roles for differences in the amplitude and phase spectra of the various melodies, the number of reversals of contour direction contained in the melodies, and the similarity of the rhythmic pattern of the notes of the contour. Given the existence of multiple possible factors, it is of interest to determine the relative contributions of each of these factors to perceived contour similarity. As a first step towards this assessment, it is important to determine how related the predictors are to one another. Accordingly, all of the predictors described earlier were intercorrelated; the resulting half-matrix of correlations appears in Table 2. As might be anticipated, there were

TABLE 2. Similarity Predictor Intercorrelation Matrix

			Pitch Code				Oscillation Model				Surface Correlation		
			Weighted		Unweighted		Reversals		Interval Size		Pitch		Rhythm
			Amp	Phs	Amp	Phs	Sum	Mean	Sum	Mean	Unw	Wgt	
Pitch Code	Weighted	Amp	—	.17	.69**	.11	.20	.18	.09	.25*	−.23*	−.46**	−.08
		Phs		—	.15	.32**	.09	.25*	−.07	−.10	−.08	−.24[A]	−.08
	Unweighted	Amp			—	−.12	.08	−.06	.14	.16	−.21	−.49**	−.01
		Phs				—	.20	.12	.03	−.13	−.10	−.23[B]	−.21
Oscillation Model	Reversals	Mean					—	.36**	.43**	−.08	−.31*	−.10	−.22
		Sum						—	.28*	−.14	−.05	−.04	−.05
	Interval Size	Mean							—	.08	−.30*	−.05	.07
		Total								—	−.08	−.20	−.25*
Surface Correlation	Pitch	Unw									—	.15	−.03
		Wgt										—	.11
	Rhythm												—

*$p < .05$ A $p = .05$
**$p < .01$ B $p = .06$

significant correlations between predictors that were conceptually related. For instance, the unweighted and weighted difference scores for both amplitude and phase spectra were significantly correlated, as were a number of the oscillation measures (e.g., summed and mean reversals, summed interval size and summed reversals, mean interval size and summed reversals).

Of more interest, however, are the relations between families of predictors, such as the Fourier analysis predictors and the oscillation measures, or the Fourier analysis predictors and the surface correlation predictors. Interestingly, and somewhat reassuringly, by and large the different families of predictors were unrelated, with only a few notable exceptions. For instance, contour similarity determined on the basis of surface correlation similarity for the weighted pitch codes appears to be related to the Fourier analysis measures of the unweighted and weighted pitch codes. In the same vein, similarity values based on the surface correlations of the unweighted pitch codes were correlated with two of the oscillation model values— summed reversals and summed interval size. Although this observed non-independence between some factors is potentially worrisome, it must be remembered that the surface correlations of the pitch codes were unrelated to listeners' similarity judgments. Accordingly, the interrelations between these factors plays no role in the assessment of the relative strengths of the significant predictors.

Given these analyses, it is then viable to predict perceived similarity ratings from a set of the previously described factors. Specifically, the amplitude and phase spectra difference scores for the weighted pitch code, the summed contour reversals, and the rhythm code surface correlations were found to significantly predict listeners' similarity ratings, $R = .58$, $p < .001$. Of the four variables entered into this analysis, the rhythm surface code correlation added significantly, $B = 1.79$, $\beta = .45$, $p < .01$, and the phase spectra difference scores were marginally significant, $B = -.49$, $\beta = -.18$, $p = .09$. Neither the amplitude spectra difference scores, $B = -.07$, $\beta = -.12$, *n.s.*, nor the summed reversals, $B = -.05$, $\beta = -.13$, *n.s.*, added significantly.

Discussion

This study demonstrated that listeners' similarity judgments for pairs of melodic contours were related to a set of theoretical predictors based on a number of different factors, including the relative strengths of different frequency cyclic patterns present in the contours, the phase relations between these cyclic components, the degree of contour oscillations, and the degree of overlap between the rhythmic pattern of the contours. The success of at least two of these factors, namely the role of amplitude spectra information and contour oscillations, is not surprising in that both converge with previous findings (Eerola & Bregman, 2007; Eerola et al., 2006; Prince et al., 2009a; Schmuckler, 1999). As such, not only do these data replicate these earlier studies, they also extend the efficacy of these models to direct similarity judgments of variable length, rhythmically complex, melodic contours.

There are, however, a number of novel findings in this study. First is the finding that employing rhythmically weighted pitch codes produced better predictive power than rhythmically unweighted pitch codes; this finding will be explored in the General Discussion. Second, there is the finding that the phase relations for the weighted pitch codes also predicted contour similarity. The fact that phase information predicted similarity is not unique—Schmuckler (1999) did, after all, observe a role for phase in similarity of simple tonal melodies written to contain explicit phase relations, and Prince et al. (2009a) also found an intermittent role for phase spectra similarity. Third, and most intriguingly, is the finding that the overlap of the rhythmic pattern of the contour also played a role in contour similarity. The most straightforward interpretation of this finding is that, despite being asked to judge similarity solely on the basis of the pitch contour itself, listeners nevertheless incorporated rhythm into their similarity judgments. In some ways this is not all that surprising a result. Listeners were not explicitly instructed to ignore rhythm in their judgments (although the experimenter did focus their attention on pitch), and given that attending to rhythm is a natural part of listening to music, there is no reason why listeners should have ignored this information. As such, it seems quite reasonable that rhythm would play some role in contour perception, and be a factor in perceived similarity.

Most importantly, and taken in conjunction with the findings of the multiple regression analysis, is the suggestion that perceived contour similarity is influenced by two factors—one based on the degree of overlap among the cyclic pitch relations contained in melodic contours, and a second based on the degree of overlap of the rhythmic pattern of the melodic contours. The possibility of these two factors at work in contour perception has a number of implications. Of primary significance is the suggestion that if one were to present melodic contours devoid of rhythmic variation, then perceived contour similarity would still be related to similarity based on amplitude spectra information but not on rhythmic overlap. Exploring this question was the goal of Experiment 2.

Experiment 2: Similarity of Equitemporal Folk Melodies

This study tested the idea that perceived contour similarity would remain largely consistent when listeners are asked to judge equitemporal, as opposed to rhythmically complex, melodic contours. To examine this question, the same set of melodies employed in the previous experiment were modified by removing all temporal deviations in note lengths, and having listeners once again provide similarity ratings for pairs of these contours.

Method

PARTICIPANTS

The sample of listeners consisted of 18 undergraduate students (*M* age = 19.8 yrs, *SD* = 1.6), who received either course credit in introductory psychology or were paid $7 for participating. Listeners had an average of 7.3 years (*SD* = 3.2) of experience playing an instrument or singing, spent 2.0 hours/week (*SD* = 2.7) involved in music making activities, and listened to music for an average of 12.1 hours/week (*SD* = 18.2). All listeners reported normal hearing.

EXPERIMENTAL APPARATUS, STIMULI, AND PROCEDURE

Stimuli for this study consisted of modified versions of the melodies employed in Experiment 1 (see Figure 6), in

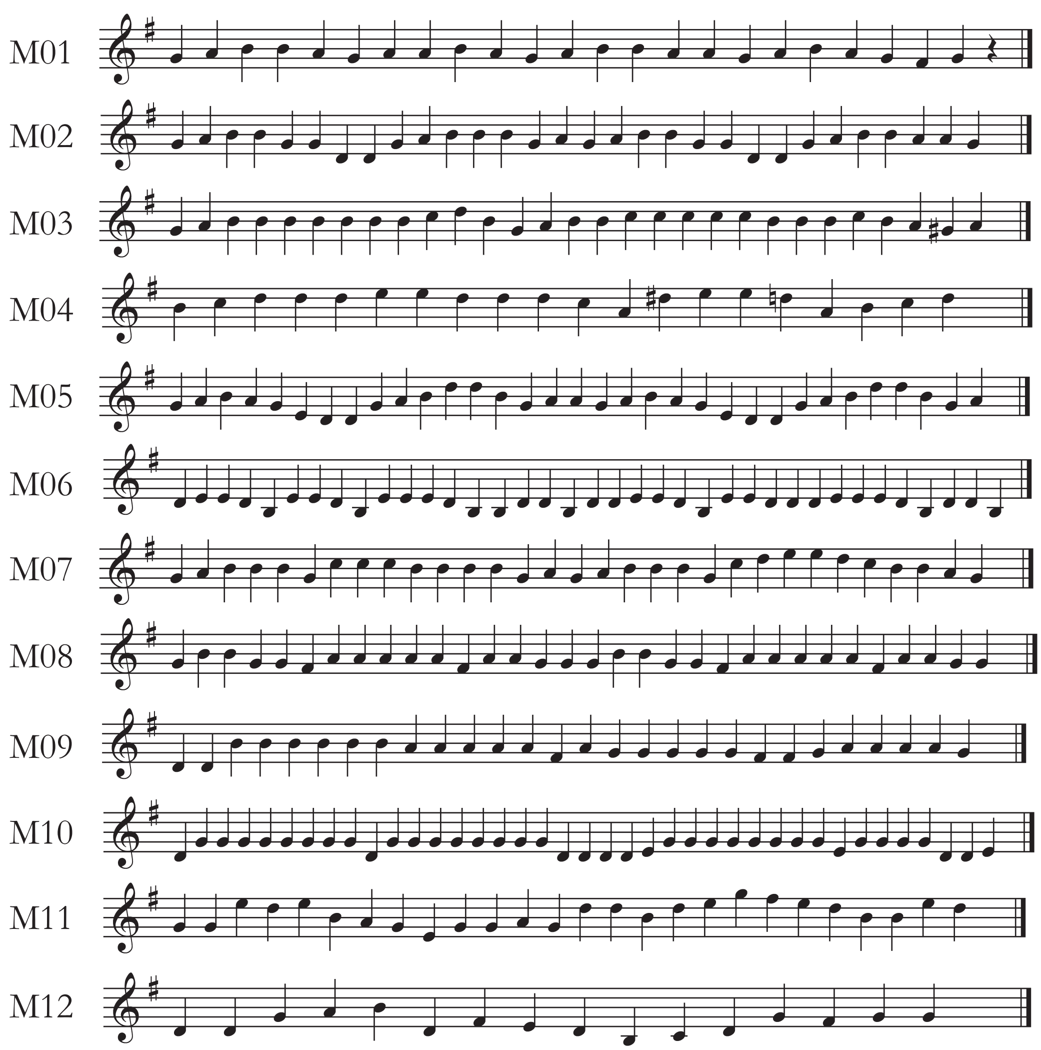


FIGURE 6. The modified, equitemporal folk melodies employed as experimental stimuli in Experiments 2 and in the equitemporal condition of Experiment 3.

which all rhythmic deviations were removed, such that each note in the melody was heard for 400 ms.[3] The remaining details of the experimental apparatus, stimuli, and procedure, were identical to those of Experiment 1.

Results

Listeners' similarity ratings were analyzed comparably to Experiment 1. Initial analyses again explored the impact of melodic ordering by correlating listeners' similarity values for the two half-matrices representing the varying melodic ordering. As in the previous study, on an individual listener basis the degree of comparability between these two half-matrices was somewhat variable, with individual listener correlations ranging from −.04 to .46 (M = .18, SD = .16). Once again, however, the two averaged half-matrices were strongly related, $r(64) = .50$, $p < .001$, indicating that similarity ratings were reasonably comparable regardless of melodic ordering.

Given this degree of correspondence on an averaged subject basis, individual listener ratings were combined across the two presentation order half-matrices, and then averaged to produce a single half matrix reflecting the averaged ratings of perceived contour similarity for all possible melodic pairs. These similarity ratings were then predicted from the same factors as in Experiment 1, as well as correlated with the similarity ratings themselves from the previous study.

Table 1 presents the correlations between listeners' similarity ratings and the various theoretical models of contour similarity, and outlines a similar pattern as observed in the previous study. For the Fourier analysis model, replicating Experiment 1, both the amplitude difference scores and the phase difference scores for the weighted pitch codes—but not the unweighted pitch codes—employed in Experiment 1 once again predicted perceived similarity. Similarly, contour reversals, this time in the form of the average number of such reversals, also correlated with similarity ratings. Moreover, the surface correlation of the rhythm codes as quantified in the previous study was again related to similarity, although this effect was now only marginally significant. Finally, the similarity ratings for this study were strongly related to those of Experiment 1, $r(64) = .71$, $p < .001$.

As in Experiment 1, multiple regression analyses were conducted to assess the relative power of the various theoretical factors in predicting perceived similarity ratings. Specifically, listeners' perceived similarity ratings were predicted from the four factors of amplitude and phase spectra difference scores for the unweighted pitch codes, the difference score for the mean number of contour reversals, and the surface correlation of the rhythm code. These variables significantly predicted listeners' ratings, $R = .44$, $p = .01$, although none of the individual variables contributed significantly. All of these variables, however, provided marginal predictive power (all p's between .10 and .20).

Discussion

Overall, this study converges with the findings of Experiment 1, demonstrating that similarity ratings for pairs of equitemporal melodies were predictable primarily from factors based on the Fourier analysis model and secondarily, from the oscillation model. Accordingly, these studies strongly support the idea that listeners are sensitive to cyclic pitch information in melodic contours, and use such information in determining the perceived similarity of melodies.

This study also provides compelling support for the idea that perceived contour similarity can be driven by two factors, one based on pitch information and a second based on rhythmic patterning. This result is most easily discernible by the fact that the predictive power of rhythm pattern similarity was diminished in this study, both in terms of its simple correlation with similarity ratings and through its nonsignificant role in the multiple regression analysis. Given that this study employed equitemporal melodies that, by definition, do not contain any rhythmic variation, this result is exactly what one might anticipate. Put differently, removal of rhythmic information in the melodies themselves had the selective effect of removing the predictive power of rhythmic similarity on listeners' ratings. The fact that the simple correlation between the rhythmic code surface correlations and the equitemporal stimuli was marginally significant is at first blush nonintuitive. However, further reflection suggests that this result can be understood by realizing that the similarity ratings in this study were strongly related to the ratings of Experiment 1 (indicating the importance of pitch change information in contour perception). Accordingly, it would have been unrealistic to anticipate that rhythmic surface information

[3]The astute reader will note that the equitemporal melodies were, at least on a "beat" level, slightly sped up relative to the previous experiment. Because the original melodies contained numerous notes that were metrical subdivisions (i.e., of shorter duration) of the musical beat (which was 500 ms), equating the durations of all of the melody notes has the unfortunate consequence of increasing (dramatically in some cases) the overall length of the equitemporal versions relative to their rhythmically complex counterparts. As such, decreasing the "beat" of each note of the equitemporal melody correspondingly decreases the length of these new versions, making them closer (although not equivalent) to the duration of the original melodies.

would be totally unrelated to similarity ratings. Rather, it is more reasonable to expect a selective diminution of this effect, as opposed to a total removal.

Also converging with Experiment 1, this study confirmed the finding of a relation between the phase spectra difference scores for the weighted pitch codes and listeners' similarity ratings. The replication of this finding is provocative in that it does suggest that listeners are, on some level, attuned to the relatively timing within the melody of the up-down (or down-up) patterns of pitch change, and do use such information in their contour assessments.

Together, Experiments 1 and 2 also support Schmuckler's (1999) claim that the calculation of a derived similarity measure based on contour complexity judgments is a viable approach for quantifying the perceived similarity of melodic contours. Such support is provided by the fact that the same model of contour (the amplitude and phase spectra values produced by a Fourier analysis) remained effective in predicting similarity judgments regardless of whether such judgments were gathered directly (as in the current work) or indirectly (as in Schmuckler, 1999). Fortunately, the current context allows for a more explicit comparison of employing direct versus derived similarity measures procedures. Specifically, it is possible to present the folk melodies used in the previous two studies to listeners, and ask them to provide contour complexity judgments, as in Schmuckler (1999) and Eerola et al. (2006). These complexity judgments can then be used to calculate a derived similarity measure, with these derived similarities then compared to the current models of contour similarity as well as the direct similarity judgments of Experiments 1 and 2 themselves.

Experiment 3: Contour Similarity From Contour Complexity Judgments

The primary goal of Experiment 3 was to determine whether similarity judgments based on contour complexity ratings are comparable to direct similarity judgments. Towards this end, listeners heard both the rhythmically diverse and equitemporal folk melodies of the previous studies and provided contour complexity judgments for these melodies. These complexity judgments were then used to derive similarity ratings between the melodies, and then predicted from the various theoretical models of described earlier.

Method

PARTICIPANTS

The sample of listeners consisted of 25 undergraduate students (*M* age = 20.8 years, *SD* = 2.0), who received either course credit in introductory psychology for participating or who volunteered their time. Although these listeners were not initially screened for musical training, they did evince a range of musical experience. Specifically, listeners had an average of of 9.7 years (*SD* = 4.7) of experience on an instrument or voice, spent 3.0 hours/week (*SD* = 3.4) involved in music making activities, and listened to music for an average of 17.4 hours/week (*SD* = 24.4). All listeners reported normal hearing.

EXPERIMENTAL APPARATUS, STIMULI, AND PROCEDURE

Stimuli consisted of the rhythmically complex and equitemporal versions of the folk melodies employed in Experiments 1 and 2, and made use of the same apparatus as in the previous two studies. On each trial listeners heard a melody, after which the computer prompted for a rating of the complexity of the melody's contour on a "1" ("not at all complex") to "7" ("very complex") scale. After listeners typed in their response, the computer paused for 400 ms, and the next trial began automatically. Overall, listeners completed two blocks of 48 randomly ordered trials, with each block containing four repetitions of the 12 rhythmically diverse melodies and the other block containing four repetitions of the 12 equitemporal melodies, with the order of these two blocks (rhythmic then equitemporal versus equitemporal then rhythmic) counterbalanced across listeners. Unfortunately, a programming error resulted in the loss of the data for the first block of trials. Fortunately, because block order was counterbalanced, the end result of this error was that a variable that was initially intended to be a within-subjects factor (rhythmic versus equitemporal melody) was transformed into a between-subjects variable (12 listeners contributed ratings for the rhythmic stimuli and 13 listeners produced ratings for the equitemporal stimuli). After completing all of the trials, listeners were debriefed as to the purposes of the experiment. The entire experimental session lasted approximately 30 minutes.

Results

Complexity ratings for each listener for the rhythmic and equitemporal melodies were averaged across the four repetitions and analyzed in a two-way analysis of variance (ANOVA), with the within subject factor of *melody* (melody 1, melody 2 . . . melody 12) and the between-subjects factor of *rhythm type* (rhythmic versus equitemporal). Although this analysis does not investigate the primary experimental hypotheses (i.e., whether or not derived similarity ratings based on contour complexity ratings of individual melodies are comparable to direct similarity ratings of melodic contours), it does assess whether the complexity ratings for the different melodies varied across rhythm type. This analysis revealed a main effect for melody, $F(11, 253) = 5.69$, $MSE = 2.72$, $p < .001$, but no effect for rhythm type, $F(1, 23) = 0.03$, $MSE = 5.83$, *n.s.*, and no interaction between the two factors, $F(11, 253) = 0.64$, $MSE = 2.72$, *n.s.*

Following Schmuckler (1999), similarity ratings based on these complexity judgments were calculated by creating a single complexity vector for each melody, produced by aggregating ratings across the listeners, and then correlating these complexity vectors. This procedure resulted in two sets of 66 correlations, representing derived contour similarity between all pairs of rhythmic and equitemporal melodies. Overall, the derived similarity scores from these two sets of melodies were strongly related, $r(64) = .60$, $p < .001$; accordingly, a third set of derived similarity scores were created by averaging the rhythmic and equitemporal scores. All three of these contour similarity measures were then predicted from the predictors employed in the previous experiments.

Table 1 presents the results of these predictions. Overall, the observed pattern of correlations strongly converges with the previous studies. Replicating both earlier experiments, derived similarity ratings were predictable from both amplitude and phase spectra difference scores for the weighted, but not the unweighted, pitch codes. The oscillation model also predicted listeners' ratings, with mean interval size correlating with similarity values. Also similar to the previous experiments, the surface correlation of the rhythmic code predicted contour similarity, with this relation diminished (i.e., marginally significant) for the equitemporal melodies. Finally, the derived similarity ratings for the rhythmic and equitemporal melodies of this experiment were related to the similarity ratings of Experiments 1 and 2, with $r(64) = .51$, $p < .001$, for the rhythmic stimuli, and $r(64) = .64$, $p < .001$, for the equitemporal stimuli.

Given the presence of multiple correlates for these similarity values, multiple regression analyses were performed to assess the relative strengths of these predictors. Specifically, the three sets of similarity scores were predicted from the amplitude and phase spectra difference scores for the weighted pitch codes, the mean interval size difference scores, and the surface correlations of the rhythmic codes. For the rhythmic stimuli, derived similarity scores were significantly predicted from these four variables, $R = .55$, $p < .001$, with two factors contributing significantly – the amplitude spectra difference scores, $B = -.08$, $\beta = -.29$, $p < .05$, and the mean interval size difference scores, $B = -.12$, $\beta = -.24$, $p < .05$. For the equitemporal stimuli, derived similarity scores were also predictable from these four variables, $R = .67$, $p < .001$, with three of the four factors contributing to this prediction – amplitude spectra difference scores, $B = -.08$, $\beta = -.27$, $p = .01$, phase spectra difference scores, $B = -.31$, $\beta = -.24$, $p < .05$, and mean interval size, $B = -.25$, $\beta = -.49$, $p < .001$. Finally, the averaged rhythmic and equitemporal similarity values were also predictable from these factors, $R = .676$, with three factors contributing significantly—amplitude spectra difference scores, $B = -.08$, $\beta = -.31$, $p < .005$, phase spectra difference scores, $B = -.26$, $\beta = -.23$, $p < .05$, and mean interval size, $B = -.19$, $\beta = -.42$, $p < .001$.

Discussion

The primary result of this experiment was that contour complexity ratings of individual melodies produced similarity values that are comparable to direct similarity ratings of pairs of melodic materials. Most notably, the degree of overlap in the amplitude spectra of melodies produced by a Fourier analysis of the pitch content of these melodies provided a reasonable prediction of perceived similarity, regardless of whether such similarity was assessed through direct or indirect techniques. Such a finding provides strong support for the notion that cyclic pitch information is an important component of listeners' percepts of contour, and furthermore, that such cyclic information can be assessed via time series techniques.

In keeping with the earlier experiments, this study also observed an impact for the overlap of phase spectra information in predicting similarity, as well as for one of the parameters of the oscillation model. The fact that both phase and the oscillation models once again correlated with derived similarity is particularly noteworthy in that this was the first context in which all these factors contributed significantly to perceived similarity, as assessed via multiple regression. The significance of these results will be more fully explored in the General Discussion.

This study did reveal important differences as a function of direct versus indirect similarity measures. One such distinction is the fact that correlations between the direct similarity ratings of Experiments 1 and 2 and the derived similarity ratings of Experiment 3 were not as strong as the correlation between different samples of direct similarity judgments (i.e., Experiments 1 and 2). As such, derived similarity measures are clearly not fully interchangeable with direct similarity assessments. One possible reason for this difference might involve the means of deriving similarity scores from complexity ratings. Because individual listeners' complexity ratings were aggregated for each melody, with these aggregate melodic complexity scores then correlated, the resulting similarity scores are strongly susceptible to individual differences in the use of the subjective rating scale. Although gathering multiple repetitions for each stimulus induces some measure of stability on an individual subject level, this analysis nevertheless can be strongly influenced by outlying values produced by particular listeners. In contrast, both Experiments 1 and 2 employed similarity ratings averaged over 16 (Experiment 1) and 18 (Experiment 2) listeners, and thus likely contains a far greater degree of stability.

A second explanation for this difference, and certainly not a mutually exclusive one, arises because similarity measures from Experiments 1 and 2 were (presumably) produced by direct comparisons of the melodic contour, whereas similarity scores for Experiment 3 were filtered through the task of rating melodic complexity. As discussed by Schmuckler (1999), musical, or in this case melodic, complexity is, in and of itself, a psychologically intricate concept—one that has been a topic of a great deal of research in its own right (Arkes, Rettig, & Scougale, 1986; Conley, 1981; Eerola et al., 2006; Konečni, 1982; Rohner, 1985; Williams, 2004). Accordingly, the divergence between these sets of ratings might highlight those factors that are intrinsic to the notion of melodic complexity, but have little to do with contour structure per se.

In this vein, it is interesting to note that one important difference between the direct and indirect similarity measures was the lack of a contribution by the surface correlation of the rhythm codes to the multiple regression model predicting derived similarity. One possibility is that, relative to the pitch variation, the rhythmic variation of these melodies was fairly simple; hence, the role of rhythm in driving complexity was reduced. Subsequent work might profitably investigate this idea, systematically varying the degree of variability in pitch and rhythm across a set of melodies, and looking at the impact of this manipulation on both complexity ratings as well as predictions of derived contour similarity.

General Discussion

Three experiments examined the impact of pitch and rhythm factors on listeners' direct and indirect similarity judgments of melodic pairs. Within each of these categories, there was evidence that some of the highlighted factors played a role in perceived contour similarity. The relative degree of success for these factors has a variety of implications for our understanding of melodic contour processing.

One of the primary findings of this work is its demonstration that the degree of convergence in the cyclic pitch content between melodies, as indexed by the amplitude spectra produced by a Fourier analysis of a representation of the pitch contour of the melody, was correlated with perceived contour similarity. This result replicates the model initially proposed by Schmuckler (1999), and extends these ideas by demonstrating that the Fourier analysis model is applicable to longer, more naturalistic rhythmically complex melodies. Moreover, these findings converge with related research (e.g., Eerola & Bregman, 2007; Eerola et al., 2006; Prince et al., 2009a) in providing compelling evidence for the viability of time series analyses as tools for understanding melodic structure.

One surprising finding in this work was the recurrent effect for phase information in predicting contour similarity. As already discussed, Schmuckler (1999) found that phase was inconsistent in predicting contour similarity, with this factor playing a role for simple tonal melodies that were explicitly written to contain certain phase relations, but not for more complex 12-tone rows. Similarly, Prince et al. (2009a) also found an inconsistency in phase predictions, with the similarity of short (about 17 notes) but not long (about 35 note) melodies predicted by phase relations. In contrast, all three experiments in this project demonstrated a correlation between phase difference scores and similarity ratings. Clearly, it seems that there is a role for phase relations in contour similarity, with this role (presumably) varying with specific parameters of a melody. As just mentioned, although specific parameters are as yet unclear, based on Prince et al.'s results, one possibility might involve melodic length. One goal of future work will be to address this finding more thoroughly, providing a systematic quantification and manipulation of melodic structure and looking at its impact on phase predictiveness.

One of the most important and novel findings of this work was the fact that for both the amplitude and phase spectra difference scores, employing a weighted versus unweighted pitch code had a large effect on predicting contour similarity. In understanding this difference it must be remembered that one important consequence of weighting pitch events by their individual durations is that it changes the framework under which the resulting frequency information is interpreted. Specifically, unweighted pitch codes capture cyclic information in a time frame employing the melody itself as the unit of analysis. In contrast, weighted pitch codes capture cyclic information in a more absolute time frame, given that durations are coded (albeit in arbitrary units).

Accordingly, these findings convincingly demonstrate that the cyclic information involved in the contours is perceived by listeners in a set time frame. The importance of timing information for phase differences is relatively easy to understand in that phase, by definition, is sensitive to timing aspects of the cyclic information captured in a Fourier analysis. The fact that the amplitude spectra information also required this timing information is somewhat more intriguing, in that this information captures the strength of the up-down-up patterns in a contour, but not their timing (e.g., the difference between up-down-up and down-up-down). The most obvious explanation here, though, is that it is not as much the time frame that is important, but rather, that weighting captures important rhythmic variation in when pitch changes occur; this aspect will be returned to shortly.

Accordingly, the inclusion versus noninclusion of timing information in the pitch code will have little impact on amplitude spectra information, an idea supported by the fact that differences scores based on amplitude spectra information were highly similar (see Table 2). Duration information, however, will influence phase spectra relations in its potential to change the relative timing relations of the cyclic patterns highlighted by a Fourier analysis; this idea is supported by the relatively weak (albeit significant) correlation between phase spectra difference scores for the weighted and unweighted pitch codes. As such, the finding that the weighted pitch spectra information predicted listeners' ratings suggests that listeners did make use of the relative timing of cyclic contour information. Put differently, and somewhat more simply, the difference between an up-down-up and a down-up-down melodic pattern is an important aspect of perceived contour relations, even though they both contain the same strength of their cyclic information.

Given the success of the Fourier analysis model in this and other work, it is important to consider the limitations of this approach as a general model of melodic contour. In a very basic sense, the application of time series analyses to quantifications of melodic structure raises some troubling questions. For one, there are any number of problematic issues related to the appropriateness of applying this mathematical technique to a discrete series (see Schmuckler, 1999, for discussion of this concern). To put it simply, the use of Fourier techniques in this context is admittedly a nonstandard application. Because signal processing techniques such as Fourier analysis assume that the signal to be analyzed is theoretically temporally infinite, analysis of truncated signals potentially produce multiple artifacts. To counteract many of these artifacts, given that any analyzable signal is in practice finite, various techniques such as windowing and detrending of the signal are routinely employed, although it should be noted that such techniques introduce their own artifacts. It is important to remember, though, that the primary consequence of any of these potential problems is that the quantification of the cyclical information contained in the signal becomes noisy and distorted. On a practical basis, the end result of such distortion would be to mask, or possibly eliminate, the ability to uncover an effect of this signal processing information vis a vis perceived contour similarity. Accordingly, the fact that these studies, as well as previous work (Prince et al., 2009a; Schmuckler, 1999), still observed these relations makes these findings all the more striking, given that they likely underestimate the predictive power of such factors.

A second, more conceptual concern arising out of the use of this procedure is that presenting Fourier analysis as a model of melodic structure raises, at least on some level, the question of whether this is literally a model of the psychological processing of melodic contours. In other words, does the success of the Fourier analysis approach imply that listeners are literally conducting Fourier transforms of melodic contours upon hearing them, and using the resulting output as a means of understanding these melodies?

Although it might be tempting to answer this question affirmatively, it does seem unlikely that there is a literal Fourier transform occurring upon hearing a melody. Instead, it seems more probable that when listeners hear a melody they simply take note of the general up-down (or down-up) pattern of this melody, and how quickly such patterns occur. Few such cycles over the course of the melody will be characterized by an amplitude spectra with greater power in its lower harmonics with particular phase values, whereas a large number of up-down cycles will result in a greater dispersion of power to the upper harmonics of the amplitude spectra. Such rough characterizations of the differing frequency information could then be used in explicit contour comparisons, or in individual contour complexity judgments. As an aside, it is worth noting that the idea of simultaneously noting the prevalence of low versus high frequency variation is conceptually akin to a model of contour processing proposed by Gilden, Schmuckler, and Clayton (1993). These authors argued that the perception and discrimination of visual fractal contours could be modeled by simultaneously noting the presence and strength of smooth, low frequency information relative to the availability and strength of high frequency, point-to-point fluctuations. Such a dual component model seems especially appropriate for these findings, although future work needs to address this possibility more explicitly before such an argument could be accepted.

Moving beyond the results pertaining to the Fourier analysis model, another noteworthy finding of this study is the failure of the surface correlation of the raw pitch codes to predict perceived similarity. This result is important in that it undermines Quinn's (1999) music-theoretic model of contour similarity (Shmulevich, 2004), in which the surface correlation of melodic contours was predictive of contour similarity, and contradicts the results of Prince et al. (2009a), who did find an effect for surface correlation. What accounts for this divergence in findings? With regard to Prince et al. (2009a), the most obvious difference between these projects is that the current study focused on auditory contour perception, whereas Prince and colleagues examined crossmodal perception. The idea that crossmodal comparisons might drive observers to focus more on surface parameters is intriguing, and worthy of study in its own right. Speculatively, one implication of such a result is that it suggests that such surface parameters might represent an amodal invariant of such contours. Hence, in crossmodal contexts such aspects are critical; in contrast, in a unimodal context attention can be given to alternative sources of structure.

As for the findings of this work and Quinn (1999), the most obvious distinction between these two sets of studies involves the length of the stimuli employed, with Quinn's work using melodies that were shorter than Schmuckler (1999) and the current project. As has been suggested elsewhere (Schmuckler, 2004), the length of the stimuli is important in that shorter melodies are not especially amenable to time series analyses, given that they provide amplitude and phase spectra with few components. In contrast, shorter melodies are particularly appropriate for analyses of surface patterns, given that listeners can easily and accurately retain such surface information in short term memory. As melody length increases, however, the relative importance of these two contour characterizations will reverse. Because surface characteristics of melodies will become increasingly difficult to remember, listeners may then make increasing use of more generalized information, such as the relative strengths of low and high frequency fluctuations. Accordingly, there will be a corresponding increase in the use of Fourier spectra information. If true, then as suggested by Schmuckler (2004), the relative importance of the two different characterizations will vary systematically with melody length, as well as potentially providing an arena of overlap in which both models are applicable to melodies of the same length. Current work is exploring this possibility.

Another novel result of these studies involves their demonstration of a crucial role for rhythm in contour perception. This influence was apparent in two ways. First, there was the result that the Fourier analysis information based on the rhythmically weighted, but not the unweighted, pitch codes consistently predicted listeners' perceived similarity. Accordingly, it appears important that the pitch codes specifically contain rhythmic variation, suggesting an obvious role for this parameter in contour similarity. Second, there are the results regarding the predictive power of an explicit copy of the rhythmic information. In this case, it is noteworthy that the rhythm code surface correlations predicted similarity ratings in Experiment 1, and for the rhythmic stimuli of Experiment 3, but decreased in predictive power in Experiment 2 and for the equitemporal stimuli of Experiment 3. The fact that the influence of this variable diminished in Experiment 2 is not surprising – these contours were, after all, equitemporal. The fact that rhythmic influence disappeared in its predictive power for both sets of stimuli in Experiment 3 (based on the multiple regression analyses) is somewhat surprising, and as has already been discussed, might indicate that because of the relative simple rhythmic pattern of these melodies, it did not influence judgments of melodic complexity. Overall, however, these findings do suggest an important general role of rhythm in contour perception.

These studies also assessed a third family of characterizations of contour structure, based on the degree of oscillation present in the melodies. Interestingly, although there was some support for this type of factor in these studies, it was clearly not overwhelming. On the one hand, all three experiments demonstrated a role for contour oscillations, with contour reversal information important in Experiments 1 and 2, and interval size information playing a role in Experiment 3. These findings do replicate Schmuckler's (1999) finding of an impact for such information, and converge with Eerola and Bregman (2007), who observed an effect of "interval content," which incorporates a measure of the mean interval sizes contained with a contour, on contour similarity.

On the other hand, contour oscillation did not contribute significantly to the multiple regression predictions of Experiments 1 or 2, although they did play a role in Experiment 3. Accordingly, the importance of this factor is inconsistent, at best, with this variation undermining the power of this variable as a critical component of contour descriptions. Again, there are precedents for such variability. In Schmuckler (1999), for instance, although interval size information was consistently related to contour similarity (contour reversals were not uniformly predictive of these ratings across experiments), this factor failed to contribute significantly in multiple regression contexts. Unfortunately, an explanation for why this model is so variable in its predictive power is not immediately forthcoming. Clearly, future work should continue to explore this factor, attempting to delineate the contexts in which this factor asserts itself.

Finally, it also should be highlighted that there is at least one family of factors that are conspicuous in their absence in the current studies as predictors of contour similarity. This family of factors can best be thought of as arising from "structural aspects" of musical passages, and includes aspects such as the organization of pitch and temporal information into complex tonal and metric hierarchies of stability. Given the great deal of attention paid to tonal (e.g., Krumhansl, 1990, 1991, 2000a; Tillman et al., 2000) and metrical (e.g., Jones, 1976; Lerdahl & Jackendoff, 1983; Palmer & Krumhansl, 1990; Povel, 1981, 1984) hierarchies in music cognition generally, and in melodic perception specifically (e.g., Boltz, 1989a, 1989b; Dowling, 1978; Jones & Boltz, 1989; Schmuckler, 2004, 2009), it is clearly surprising that such hierarchal information has not been included in this analysis of contour perception.

Once again, however, evidence for such factors is mixed. For example, Eerola and colleagues (Eerola & Bregman, 2007; Eerola et al., 2006) recently demonstrated that such factors play a role in both contour complexity judgments and contour similarity, and tonality (in the form of pitch distributional information) has been implicated in the perception of melodic similarity as well (Eerola et al., 2001). Similarly, van Egmond, Povel, and Maris (1996) found that the perceived similarity of melodies and their transpositions was driven by the frequency distance, as well as the tonal relatedness, of the original melody and its transposition. More generally, both contour and tonal information have been highlighted as fundamental factors in the perceptual organization of, and subsequent memory for, an array of melodic materials (Dowling, 1978; Schmuckler, 2009). As such, one might expect that tonal information would play a role in contour similarity.

Regardless of this expectation, though, explicit tests for a role of tonal and metric hierarchy similarity in predicting contour similarity have often failed to observe an effect. For instance, Schmuckler (1999) found no influence of tonal similarity on contour similarity. Moreover, a series of analyses examined whether the observed contour similarity ratings were predicted by the tonal and metric relatedness of the melodies, using similarity measures based on the Krumhansl-Schmuckler key-finding algorithm (Krumhansl, 1990) employed by Schmuckler and Tomovski (2005), and metric similarity measures derived from the metric hierarchy studies of Palmer and Krumhansl (1990) and employed in Schmuckler (1990). Although not formally presented in this paper, the Appendix presents the correlations between the similarity ratings of these studies and a family of tonal and metric hierarchy similarity measures. Inspection of these correlations reveals only marginal predictive power for either Experiments 1 or 2 based on tonal and metric hierarchy similarity. For Experiment 3, some of the metric hierarchy factors were predictive of the rhythmic and averaged stimuli, along with a single significant correlation for one of the tonal hierarchy measures. Accordingly, the power of tonal and metric hierarchy information to predict contour similarity in this context is, at best, weak, and more conservatively, neglible.

What accounts for this lack of any strong effect for the otherwise critical factors of tonal and metric hierarchies? For tonality, the obvious explanation is that both the current studies and Schmuckler (1999) explicitly reduced the importance of this factor by playing all melodies in highly related keys or using atonal melodies. The reason tonality was reduced in these projects was, essentially, pragmatic. In Western music, pitch structure and tonality plays a predominant role in listeners' apprehension of music, and sometimes dominates listeners' perceptions of musical passages (Prince, Schmuckler, & Thompson, 2009b; Prince, Thompson, & Schmuckler, 2009). As such, tonal effects were negated so as to provide the best possible context in which to assess the impact of the various contour specific models. Clearly, however, future work should integrate these two factors into a more encompassing structural model of melodic information, a fact that has been recognized already (Dowling, 1978; Schmuckler, 2004, 2009).

As for the failure of metric hierarchy information, this result is admittedly curious, and not open to an easy explanation. One possibility is that although metric hierarchy information was available in these studies, this information was not as salient in these studies as the pitch information, with listeners simply more attuned to pitch changes than temporal changes. As just mentioned, the idea of a predominant status for pitch information for (at least Western) listeners has been supported by other findings (Hannon & Trehub, 2005; Prince et al., 2009b; Prince et al., 2009; Thompson, 1994; Thompson et al., 2001). Prince et al. (2009), for instance, looked at the relation between tonal and metric hierarchy information in a series of experiments, and found that pitch information predominated in listeners' percepts of these passages. Moreover, listeners' judgments of these passages continued to show an influence of pitch information even when participants were explicitly instructed to ignore this factor, and speeded classifications of tonal and metric events revealed asymmetric classifications, with pitch content influencing judgments of metric information, but not the reverse. Taken together, these findings are consistent with the idea that pitch information is highly salient in melodic processing, and dominates listeners' apprehension of musical information.

Of course, the preceding discussion of additional factors that might predict contour similarity raises one of the thornier concerns with this project, namely, the fact that although significant, the current models in general, and the Fourier analysis model in particular, did not on an absolute basis account for an especially large percent of the variance of listeners' perceived contour similarity. Although undeniably true, there is an array of responses that could be given to this criticism. First, and harkening back to a point made earlier, given the mathematical limitations inherent in the application of Fourier analysis to the melodies examined here, it is likely that this estimation of the importance of cyclical information is indeed conservative. Second, and stemming from a more practical basis, the level of predictive power of this study, although modest, has been replicated across experiments in the current project, and is consistent with previous work on auditory and crossmodal contour similarity (Eerola & Bregman, 2007; Eerola et al., 2006; Eerola et al., 2001; Prince et al., 2009a; Schmuckler, 1999). Third, and most critically, it has been assumed throughout this work that the quantification of cyclical information will not capture all of the variance in perceived contour similarity. As stated repeatedly, because melodies are complex, multidimensional auditory objects, an adequate characterization of their structure, and of melodic similarity, will rely on multiple simultaneously operative models. Accordingly, it is not worrisome that the range of identified factors contributing to contour similarity in this project is not exhaustive.

Ironically, admitting this limitation leads to the obvious next question of what additional properties might be of interest in providing a more comprehensive characterization of melodic contour. In this regard, some candidates arise through a closer examination of the actual melodies employed in this study. Intriguingly, such an exercise reveals a number of characteristics that may be important. For instance, various subsets of the melodies share a common melodic pattern at the beginning: M01 and M11 contain a major 6th interval at the start, whereas M10 and M12 contain a perfect 4th interval, and M01, M02, M03, M04, M05, and M07 all contain stepwise motion in their first three notes. Similarly, M04 and M11, and to a lesser degree M07, all share a common tessitura, and M02, M05, and M07 all have exactly the same rhythm in their first measures, along with a significant repetition of a particular rhythmic motif (dotted 8th followed by a 16th note) more generally. Some or all of these factors could potentially play a role in percepts of contour similarity. As an aside, it is important to realize that some of these aspects are at least partially captured by various measures employed in this project. The similarity of the opening intervallic patterns will contribute (somewhat) to the degree of surface correlation between the melodies, and the similarity of the rhythmic structure is captured to an extent by the rhythm surface correlations.

Of more theoretical interest, however, is a fundamental distinction between the classes of factors tested in this project, and the types of properties just outlined. As initially suggested, one of the underlying goals of this work is to develop and test characterizations of global parameters intended to capture a given property as applied to the contour as a whole. In contrast, the present set of melodic properties encompasses primarily local contour parameters, such as the initial interval content, the occurrence of rhythmic motifs, and so on. Ironically, then, the recognition of the possible importance of such parameters reintroduces local parameters back into the discussion, as well as raising the obvious issue of the general roles of global versus local factors. Leaving aside the obvious difficulties stemming from a consistent and theoretically motivated rationale for what local parameters might be of interest (e.g., do we focus on the initial two note, or three note, or X note interval?), future work might clearly profit from combining both global and local levels of analysis. As an aside, this focus on, and recognition of, a simultaneous role for both local and global factors has proven fruitful in analyses of models of tonality (Abe & Okada, 2004; Brown & Butler, 1981; Matsunaga & Abe, 2005; Yoshino & Abe, 2004; see Schmuckler, 2009, for a review).

In sum, the current paper explored the efficacy of a series of models of contour structure in predicting listeners' direct and indirect similarity judgments. These studies converge with other work in demonstrating that perceived contour structure is multiply determined, driven by a variety of factors. Moreover, contour similarity is but a single component of the question of what drives melodic similarity, a question that has come under increasing scrutiny in recent years (e.g., Cambouropoulos, 2001; Deliège, 2001; Eerola et al., 2001; Hofmann-Engl, 2001, 2004, 2005; Lamont & Dibben, 2001; Ockelford, 2004; Toiviainen & Eerola, 2002). Understanding the factors that underlie such perceived similarity are critical in that such processes form the basis of category formation, and thus provide a compelling arena for testing the generalizability of theories of category and concept formation (e.g., Markman & Gentner, 1996, 1997; Medin, Goldstone, & Gentner, 1993; Posner & Keele, 1968; Rosch & Mervis, 1975; Tversky, 1977). Accordingly, the success of models of musical features, and their role in driving musical similarity, will hopefully shed insight into a very fundamental aspect of everyday perceptual and cognitive functioning.

Author Note

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to the author. The author would like to thank Jon Prince and three anonymous reviewers for their helpful comments on earlier versions of this paper.

Correspondence concerning this article should be addressed to Mark A. Schmuckler, Department of Psychology, University of Toronto Scarborough, 1265 Military Trail, Scarborough, ON M1C 1A4. E-MAIL: marksch@utsc.utoronto.ca

APPENDIX. Correlations Between the Tonal and Metric Hierarchy Measures and Contour Similarity Ratings for Experiments 1–3

	Experiment 1	Experiment 2	Experiment 3		
Tonal Hierarchy			Rhy	Eqi	Avg
Fourier Distance	–.14	.02	–.31*	–.05	–.13
Input Vector Correlations	.24[A]	.15	.16	.04	.11
Output Vector Correlations	.22[B]	.16	.07	–.04	.02
Metric Hierarchy					
Ratings Correlations	–.22[C]	–.05	–.41**	–.22[B]	–.35**
Count Correlations	.03	.16	.28*	.03	–.14
Input Vector Correlations	.18	–.04	.39**	.12	.28*

*$p < .05$ **$p < .01$ A $p = .06$ B $p = .07$ C $p = .08$

Descriptions of Predictors (see Krumhansl, 1990, Schmuckler, 1990, or Schmuckler & Tomovski, 2005, for further explanation).

Fourier Distance: The distance between the projections of the output vectors in Fourier space.

Input Vector Correlations: The correlation between the K-S algorithm input vectors.

Output Vector Correlations: The correlation between the K-S algorithm output vectors.

Ratings Correlations: The correlation between the metric hierarchy input vectors and the metric hierarchy ratings of Palmer and Krumhansl (1990).

Count Correlations: The correlation between the metric hierarchy input vectors and the metric hierarchy note counts of Palmer and Krumhansl (1990).

Input Vector Correlations: The correlation between the metric hierarchy input vectors.

References

ABE, J., & OKADA, A. (2004). Integration of metrical and tonality organization in melody perception. *Japanese Psychological Research, 46*, 298–307.

ADAMS, C. R. (1976). Melodic contour typology. *Ethonomusicology, 20*, 179–215.

ARKES, H. R., RETTIG, L. E., & SCOUGALE, J. D. (1986). The effect of concurrent task complexity and musical experience on preference for simple and complex music. *Psychomusicology, 6*, 51–60.

BOLTZ, M. (1989a). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 749–761.

BOLTZ, M. (1989b). Rhythm and "good endings": Effects of temporal structure on tonality judgments. *Perception and Psychophysics, 46*, 9–17.

BOLTZ, M. (1991). Some structural determinants of melody recall. *Memory and Cognition, 19*, 239–251.

BOLTZ, M. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception and Psychophysics, 53*, 585–600.

BOLTZ, M., & JONES, M. R. (1986). Does rule recursion make melodic easier to reproduce? If not, what does? *Cognitive Psychology, 18*, 389–431.

BROWN, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception, 5*, 219–249.

BROWN, H., & BUTLER, D. (1981). Diatonic trichords as minimal tonal cue-cells. *In Theory Only, 5*, 37–55.

BROWN, H., BUTLER, D., & JONES, M. R. (1994). Musical and temporal influences on key discovery. *Music Perception, 11*, 371–407.

BROWNE, R. (1981). Tonal implications of the diatonic set. *In Theory Only, 5*, 3–21.

BUTLER, D. (1989). Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception, 6*, 219–241.

CAMBOUROPOULOS, E. (2001). Melodic cue abstraction, similarity, and category formation: A formal model. *Music Perception, 18*, 347–370.

CONLEY, J. K. (1981). The psychophysical investigation of judgments of complexity in music. *Psychomusicology, 1*, 59–71.

DELIÈGE, I. (2001). Introduction: Similarity perception ↔ categorization ↔ cue abstraction. *Music Perception, 18*, 233–243.

DOWLING, W. J. (1971). Recognition of inversions of melodies and melodic contours. *Perception and Psychophysics, 9*, 348–349.

DOWLING, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*, 341–354.

DOWLING, W. J. (1994). Melodic contour in hearing and remembering melodies. In R. Aiello & J. Sloboda (Eds.), *Musical perceptions* (pp. 173–190). Oxford, UK: Oxford University Press.

DOWLING, W. J. (2001). Perception of music. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 469–498). Oxford, UK: Blackwell Publishers.

DOWLING, W. J., & FUJITANI, D., S,. (1971). Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America, 49*, 524–531.

DOWLING, W. J., & HARWOOD, D. (1986). *Music cognition*. Orlando, FL: Academic Press.

EEROLA, T., & BREGMAN, M. (2007). Melodic and contextual similarity of folk song phrases. *Musicae Scientiae, Discussion Forum 4A-2007*, 211–233.

EEROLA, T., HIMBERG, T., TOIVIAINEN, P., & LOUHIVUORI, J. (2006). Perceived complexity of western and African folk melodies by western and African listeners. *Psychology of Music, 34*, 337–371.

EEROLA, T., JÄRVINEN, T., LOUHIVUORI, J., & TOIVIAINEN, P. (2001). Statistical features and perceived similarity of folk melodies. *Music Perception, 18*, 275–296.

FRIEDMANN, M. L. (1985). A methodology for the discussion of contour: Its application to Schoenberg's music. *Journal of Music Theory, 29*, 223–248.

FRIEDMANN, M. L. (1987). My contour, their contour. *Journal of Music Theory, 31*, 268–274.

GILDEN, D. L., SCHMUCKLER, M. A., & CLAYTON, K. (1993). The perception of natural contour. *Psychological Review, 100*, 460–478.

HANNON, E. E., & TREHUB, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science, 16*, 48–55.

HOFMANN-ENGL, L. (2001, October). *Towards a cognitive model of melodic similarity.* Paper presented at the 2nd Annual Internation Conference on Music Information Retrieval, Bloomington, IN.

HOFMANN-ENGL, L. (2004). *Melodic similarity—providing a cognitive groundwork*. Available from www.chameleongroup.org.uk/research/cognitive_similarity.pdf. Accessed on September 25, 2010.

HOFMANN-ENGL, L. (2005). *An evaluation of melodic similarity models*. Available from www.chameleongroup.org.uk/research/evaluation.pdf. Accessed on September 25, 2010.

JONES, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323–345.

JONES, M. R. (1993). Dynamics of musical patterns: How do melody and rhythm fit together. In T. J. Tighe & W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm* (pp. 67–92). Hillsdale, NJ: Lawrence Elbaum Associates.

JONES, M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459–491.

JONES, M. R., BOLTZ, M., & KIDD, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception and Psychophysics, 32*, 211–218.

JONES, M. R., & PFORDRESHER, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology, 51*, 271–291.

JUSTUS, T. C., & BHARUCHA, J. J. (2002). Music perception and cognition. In S. Yantis (Ed.), *Stevens handbook of experimental psychology* (3rd ed., Vol. 1: Sensation and Perception, pp. 453–492). New York: John Wiley.

KONEČNI, V. J. (1982). Social interaction and musical preference. In D. Deutsch (Ed.), *The psychology of music* (pp. 497–516). New York: Academic Press.

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

KRUMHANSL, C. L. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology, 42*, 277–303.

KRUMHANSL, C. L. (2000a). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*, 159–179.

KRUMHANSL, C. L. (2000b). Tonality induction: A statistical approach applied cross-culturally. *Music Perception, 17*, 461–480.

KRUMHANSL, C. L., & SCHMUCKLER, M. A. (1986, August). *Key-finding in music: An algorithm based on pattern matching to tonal hierarchies.* Paper presented at the 19th annual Mathematical Psychology Meeting, Cambridge, MA.

KRUMHANSL, C. L., & TOIVIAINEN, P. (2001). Tonal cognition. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music. Annals of the New York Academy of Sciences* (Vol. 930, pp. 77–91). New York: New York Academy of Sciences.

LAMONT, A., & DIBBEN, N. (2001). Motivic structure and the perception of similarity. *Music Perception, 18*, 245–274.

LEMAN, M. (1995). A model of retroactive tone-center perception. *Music Perception, 12*, 430–471.

LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

MARKMAN, A. B., & GENTNER, D. (1996). Commonalities and differences in similarity comparisons. *Memory and Cognition, 24*, 235–249.

MARKMAN, A. B., & GENTNER, D. (1997). The effects of alignability of memory. *Psychological Science, 8*, 363–367.

Marvin, E. W. (1991). The perception of rhythm in non-tonal music: Rhythmic contours in the music of Edgard Varèse. *Music Theory Spectrum, 13*, 61–78.

Marvin, E. W. (1995). A generalization of contour theory to diverse musical spaces: Analytic applications to the music of Dallapiccola and Stockhausen. In E. W. Marvin & R. Hermann (Eds.), *Concert music, rock and jazz since 1945: Essays and analytic studies* (pp. 135–171). Rochester, NY: University of Rochester Press.

Marvin, E. W., & Laprade, P. A. (1987). Relating musical contours: Extensions of a theory for contour. *Journal of Music Theory, 31*, 225–267.

Matsunaga, R., & Abe, J. (2005). Cues for key perception of a melody: Pitch set along? *Music Perception, 23*, 153–164.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science, 19*, 1263–1271.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254–278.

Monahan, C. B. (1993). Parallels between pitch and time and how they go together. In T. J. Tighe & W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm* (pp. 121–154). Hillsdale, NJ: Lawrence Erlbaum Associates.

Morris, R. D. (1993). New directions in the theory and analysis of musical contour. *Music Theory Spectrum, 15*, 205–228.

Ockelford, A. (2004). On similarity, derivation and the cognition of musical structure. *Psychology of Music, 32*, 23–74.

Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 116–126.

Palmer, C., & Krumhansl, C. L. (1987b). Pitch and temporal contributions to musical phrases: Effects of harmony, performance timing, and familiarity. *Perception and Psychophysics, 51*, 505–518.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 728–741.

Polansky, L. (1987). *Morphological metrics: An intoduction to a theory of formal distances*. In J. Beauchamp (Ed.), *Proceedings of the International Computer Music Conference* (pp. 197–205). San Francisco, CA: Computer Music Association.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353–363.

Povel, D. J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 3–18.

Povel, D. J. (1984). A theoretical framework for rhythm perception. *Psychological Research, 45*, 315–337.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009a). Cross-modal melodic contour similarity. *Canadian Acoustics, 37*, 35–49.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009b). The effect of task and pitch structure on pitch-time interactions in music. *Memory and Cognition, 37*, 368–381.

Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1598–1617.

Quinn, I. (1999). The combinatorial model of pitch contour. *Music Perception, 16*, 439–456.

Rohner, S. J. (1985). Cognitive-emotional response to music as a function of music and cognitive complexity. *Psychomusicology, 5*, 25–38.

Rosch, E., & Mervis, C. B. (1975). Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573–605.

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

Schmuckler, M. A. (1990). The performance of global expectations. *Psychomusicology, 9*, 122–147.

Schmuckler, M. A. (1999). Testing models of melodic contour similarity. *Music Perception, 16*, 295–326.

Schmuckler, M. A. (2004). Pitch and pitch structures. In J. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 271–315). San Diego, CA: Academic Press.

Schmuckler, M. A. (2009). Components of melodic processing. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 93–106). Oxford, UK: Oxford University Press.

Schmuckler, M. A., & Gilden, D. L. (1993). Auditory perception of fractal contours. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 641–660.

Schmuckler, M. A., & Tomovski, R. (2000, November). *Tonal hierarchies and intervallic rivalries in musical key-finding.* Paper presented at the Society for Music Perception and Cognition, Toronto, Ontario.

Schmuckler, M. A., & Tomovski, R. (2005). Perceptual tests of an algorithm for musical key-finding. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1124–1149.

Shmulevich, I. (2004). A note on the pitch contour similarity index. *Journal of New Music Research, 33*, 17–18.

Smith, N. A., & Schmuckler, M. A. (2004). The perception of tonal structure through the differentiation and organization of pitches. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 268–286.

Temperley, D. (1999). What's key for key? The Krumhansl-Schmuckler key-finding algorithm reconsidered. *Music Perception, 17*, 65–100.

Temperley, D. (2001). *The cognition of basic musical structures*. Cambridge, MA: MIT Press.

Temperley, D. (2007). *Music and probability*. Cambridge, MA: MIT Press.

Thompson, W. F. (1994). Sensitivity to combinations of musical parameters—pitch with duration, and pitch pattern with durational pattern. *Perception and Psychophysics, 56*, 363–374.

Thompson, W. F., Hall, M. D., & Pressing, J. (2001). Illusory conjunctions of pitch and duration in unfamiliar tone sequences. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 128–140.

Thompson, W. F., & Sinclair, D. (1993). Pitch pattern, durational pattern, and timbre: A study of the perceptual integration of auditory qualities. *Psychomusicology, 12*, 3–21.

Tillman, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107*, 885–913.

Toiviainen, P., & Eerola, T. (2002). A computational model of melodic similarity based on multiple representations and self-organizing maps. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 236–239). Sydney, AU: Casual Productions.

Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception, 32*, 741–766.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327–352.

van Egmond, R., Povel, D.-J., & Maris, E. (1996). The influence of height and key on the perceptual similarity of transpose melodies. *Perception and Psychophysics, 58*, 1252–1259.

Williams, L. R. (2004). *The effect of musical training and musical complexity on focus of attention to melody or harmony.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

Yoshino, I., & Abe, J. (2004). Cognitive modeling of key intepretation in melodic perception. *Japanese Psychological Research, 46*, 283–297.