Effects of Emergent-Level Structure on Melodic Processing Difficulty

Frank A. Russo
Ryerson University, Toronto, Canada

William Forde Thompson
Macquarie University, Sydney, Australia

Lola L. Cuddy
Queen's University, Kingston, Canada

Four experiments assessed the influence of emergent-level structure on melodic processing difficulty. Emergent-level structure was manipulated across experiments and defined with reference to the Implication-Realization model of melodic expectancy (Narmour, 1990, 1992, 2000). Two measures of melodic processing difficulty were used to assess the influence of emergent-level structure: serial-reconstruction and cohesion ratings. In the serial-reconstruction experiment (Experiment 1), reconstruction was more efficient for melodies with simple emergent-level structure. In the cohesion experiments (Experiments 2-4), ratings were higher for melodies with simple emergent-level structure, and the advantage was generally greater in the presence of simple surface-level structure. Results indicate that emergent-level structure as defined by the model can influence melodic processing difficulty.

Received: November 22, 2014, accepted April 13, 2015.

Key words: melody, hierarchical structure, memory, cohesion, expectancy

There is a rich music-theoretic tradition of describing hierarchical structure in music (e.g., Forte, 1977; Lerdahl, 1988, 1989; Lerdahl & Jackendoff, 1983; Meyer, 1973; Narmour, 1983, 1990; Schenker, 1969, 1979). In the case of melody, this normally includes some description of the melodic surface and emergent structures that are composed of salient events (see Narmour, 1983, for a useful review). From a cognitive perspective, an important question that arises from this work is whether listeners are sensitive to these descriptions and whether they may relate in some manner to melodic processing difficulty. In other words, does ease of processing depend in some manner on emergent-level structure defined by theory? The current study investigates whether melodic processing difficulty varies with respect to music-theoretic descriptions of emergent-level structure derived from the Implication-Realization (I-R) model (Narmour, 1990, 1992).

Two leading cognitive approaches to understanding melodic complexity include information-theoretic and dynamic attending models. Information-theoretic models have focused on the development of coding systems (Cuddy, Cohen, & Mewhort, 1981; Deutsch, 1980; Leeuwenberg, 1969; Restle, 1970; Simon, 1972). A hierarchical melody with surface- and emergent-level structure can be described economically using nested codes that exploit redundancies. The codes are assumed to capture important aspects of mental representation, and empirical studies have found that melodies with shorter codes are easier to process (Boltz & Jones, 1986; Deutsch & Feroe, 1981). Dynamic attending theory has focused on the role of attention in the mental representations of melody (Jones, 1987, 1993; Jones & Boltz, 1989). Internal oscillators are presumed to entrain to levels of oscillatory structure that are defined by rhythmic and melodic accents (Large & Jones, 1999). The joint accent structure hypothesis posits that melodic processing is facilitated when rhythmic and melodic accents are in phase (Boltz & Jones, 1986; Jones, 1987; Jones & Pfordresher, 1997; Jones & Ralston, 1991). This approach has received extensive empirical support and provides more flexibility than coding systems.

The I-R model makes predictions regarding processing difficulty and provides additional flexibility regarding the instantiation of hierarchical structure (Narmour, 1990, 1991, 2000). This is accomplished by considering expectancy at different levels of structure. Two tones in sequence at any level of structure are said to be implicative, leading to bottom-up and top-down expectancies for the next note to follow. Bottom-up expectancies are Gestalt-like and proposed to be innate.[1] Top-down

[1] Narmour's innate proposal is called into question by Pearce and Wiggins (2006, 2012) work on the IDyoM model, which demonstrates that the bottom-up expectancies can be simulated by corpus-based statistical learning.

Music Perception, volume 33, issue 1, pp. 96–109, issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2015.33.1.96

expectancies are acquired through statistical learning (extra-opus) and through rule iteration (intra-opus).

Complexity at any given level of structure is thought to depend on the extent to which it fulfills expectancy (Narmour, 1990, 1992; also see Meyer, 1956, pp. 138-139). Consistent with this view, Rohrmeier and Cross (2013) recently reported that implicit learning of melodies is impeded when surface-level events frequently deny bottom-up expectancies described in the I-R model (Rohrmeier & Cross, 2013). Similarly, Loui (2012) found that statistical learning of an artificial grammar is impaired after small intervals are removed from melodies (denial of a bottom-up principle of expectancy referred to as pitch proximity). No empirical evidence exists to date regarding whether bottom-up principles of expectancy may influence melodic processing difficulty in hierarchically structured melodies.

In the current study, hierarchical structure was primarily established through the manipulation of bottom-up principles of expectancy. Note-to-note transitions within groups generally fulfilled expectancy, while note-to-note transitions between groups denied expectancy. The denials were achieved by following a small interval (three semitones or less) with a large interval (six semitones or greater). These expectancy denials occurred with temporal regularity so as to create surface-level groups of regular size with the first note of each group rising to the emergent level (see Figure 1).

The type of expectancy denial implemented in this study has been judged unexpected across a variety of contexts (Cuddy & Lunney, 1995; Krumhansl, 1995; Schellenberg, 1996, 1997). Fujioka, Trainor, Ross, Kakigi, and Pantev (2004) found that this type of denial elicits a magnetic type of mismatch negativity

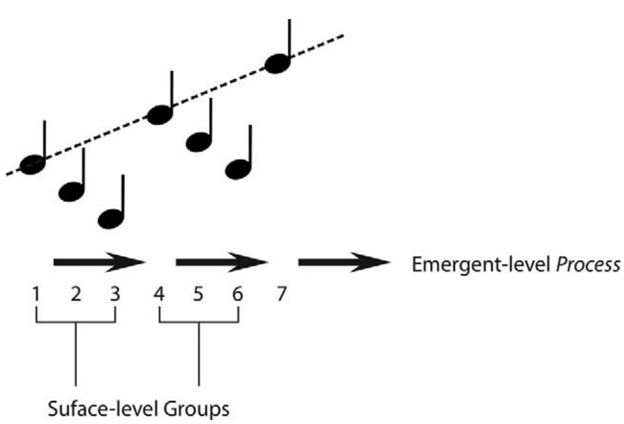

FIGURE 1. Surface-level grouping is instantiated by realizing a denial of expectancy on every fourth note. Notes 1, 4 and 7 form a highly expected emergent-level structure referred to as a *process*.

(MMNm), suggesting that it is encoded preattentively and automatically.

Emergent-level structure was further clarified using two devices. First, under certain conditions, *familiar* melodic patterns served as surface-level groups (e.g., major triad). Familiar patterns facilitate the identification of groups, which should support events rising to the emergent level. Second, under certain conditions surface-level groups were *repeated* under simple transposition. Repetition should further clarify grouping structure and draw attention to the emergent level (Deutsch & Feroe, 1981; Meyer, 1956; Margulis, 2013; Narmour, 1990, 1992, 1999, 2000).

Meyer (1973, p. 53) states that "on the hierarchic level where repetition is immediate, it [*repetition*] tends to separate events. But on the next level – where similar events are grouped together as part of some larger unit – repetition tends to create coherence." Similarly, the I-R model states that all other things being equal, the extent of repetition at the surface will influence perceived complexity and that this complexity is inversely related to attention at the emergent level (Narmour, 2000, Table 1). Hence, an immediate exact repetition of form (A^0, A^0) is more expected and will emphasize emergent-level structure more than an immediate near repetition of form (A^0, A^1), which in turn will be more expected and emphasize the emergent-level more than an immediate contrast in form (A, B). See Figure 2 for examples of different types of surface-level repetition.

Structure at the emergent level was categorically labeled as simple or complex. Simple emergent-level structure involved a sequence of small intervals moving in the same direction. This type of structure is referred to as *process* in the I-R model and is considered highly expected (Narmour, 1989, 1990, 1999). Similar structures have been described in other models as highly expected and even archetypal: *inertia* (Larson, 2012), *good continuation* (Meyer, 1956, 1973), *step inertia* (Huron, 2006; von Hippel, 2002) and *direction* (Margulis, 2005). Complex emergent-level structure also involved a sequence of small intervals but the direction of intervals varied resulting in a combination of structures that is less expected.

Two response measures were implemented to evaluate the effects of emergent-level structure on melodic processing difficulty: serial reconstruction and perceived cohesion. In the serial reconstruction task adapted from the jigsaw puzzle procedure designed by Deliège and colleagues (Deliège & Mélen, 1997; Deliège, Mélen, Stammers, & Cross, 1996; also see Tillmann, Bigand, & Madurell, 1998), a melody is presented, after which participants are given randomly arranged segments from the melody and

TABLE 1. *Demographic Information*

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Musicians				
Mean (SE) Points+	12.09 (1.03)	13.13 (0.88)	10.03 (0.59)	10.95 (0.59)
Female / Male	9 / 3	13 / 2	13 / 3	12 / 3
Mean Age (years)	21.1	19.1	21.3	20.5
Nonmusicians				
Mean (SE) Points	1.58 (0.68)	1.87 (0.34)	1.67 (0.38)	1.81 (0.69)
Female / Male	6 / 6	13 / 2	13 / 3	10 / 5
Mean Age (years)	22.0	19.3	20.2	21.2

+ All musicians continued to be musically active, whereas nonmusicians were not. One point was awarded for each year of private instruction and a half point for each year of group instruction.

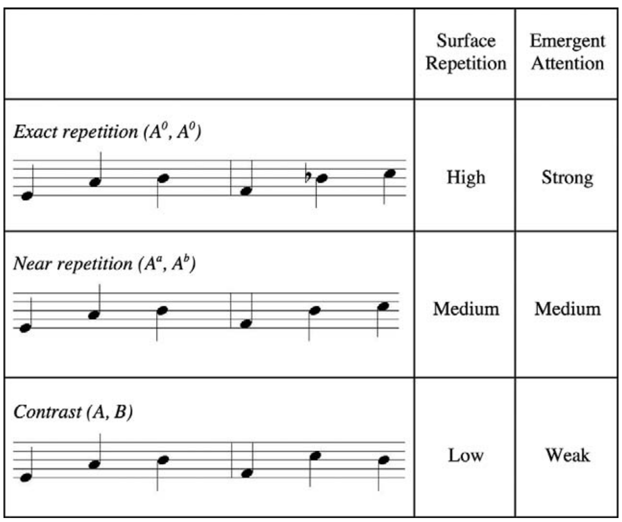

FIGURE 2. Examples of exact repetition of form (A^O, A^O), near repetition of form (A^a, A^b), and contrast in form (A, B) at the surface level.

asked to rearrange the order so as to match the original. In the cohesion task, listeners are asked to judge the perceived cohesion of melodies.

Eerola, Himberg, Toiviainen, and Louhivuori (2006) formalized a number of statistical measures to predict melodic complexity: entropy of pitch-class distribution, entropy of interval distribution, mean interval size, entropy of duration distribution, rhythmic variability, note density, tonal ambiguity, accent incoherence, contour self-similarity, and contour entropy. These measures were drawn from information-theoretic, music-theoretic, and dynamic attending approaches to melodic complexity. As we were primarily interested in the influence of emergent-level structure as defined by the I-R model, test melodies within each experiment were composed in a manner that minimized variability in these measures across levels of emergent structure (i.e., no statistically significant differences).

Experiment 1

The aim of this experiment was to assess melodic processing difficulty in hierarchically structured melodies using a serial reconstruction task. Melodies were composed to establish simple or complex emergent-level structure according to principles of the I-R model. At the surface level, melodies either repeated the same group under transposition or chained together unrelated surface-level groups. The former type of surface-level structure was referred to as simple and the latter, complex.

For melodies with simple surface-level structure, each surface-level group consisted of a major or minor triad. The group was repeated five times under transposition. For melodies with complex surface-level structure, the surface-level groups were more variable, including less familiar non-triadic sequences. We predicted main effects of simple- and emergent-level structure, as well as an interaction, whereby emergent-level differences would be enhanced when surface-level structure was simple. Ease of processing was assessed using a serial reconstruction task.

METHOD

Participants. Twenty-four undergraduate students were recruited to participate from the Queen's University community. Demographic information for each participant group (musician/nonmusician) is provided in Table 1. Participants recruited through the Introductory Psychology Participant Pool were given course credit for their participation. These participants included a mix of musicians and nonmusicians. Additional participants for the musician group were recruited using posters displayed around campus. Musicians recruited with posters were reimbursed with nominal payment.

Apparatus. Participants were individually tested in a sound-attenuated chamber. Melodies were generated from a Roland SoundCanvas tone generator, set to "Piano," under the control of a Power Mac computer running MusicShop software. Melodies were played through a single Fostex 6301 speaker monitor situated approximately one foot in front of the listener and set to a comfortable listening level. Icons were presented on the screen to represent chunks of the melody. Participants were able to rearrange the order of icons except the first using a computer mouse.

Stimuli. Eight melodies were composed to encompass all combinations of two binary factors. First, melodies possessed either simple or complex surface-level structure. Simple surface-level structure involved the repetition of a familiar melodic group (a major or minor triad). Complex surface-level structure involved a variety of melodic groups. Second, melodies possessed either simple or complex emergent-level structure. In melodies with simple emergent-level structure, the first note of each surface-level group formed a *process* at the emergent level. In melodies with complex emergent-level structure, the first note of each surface-level group formed a combination of structures at the emergent level that involved a contour change. An inverted counterpart of each melody was generated, resulting in a total of eight test melodies (ascending and descending variants of each binary combination of factors).

The music notation for each test melody is provided in Figure 3. All test melodies shared the following characteristics: 15 tones (9 tones occurring once and 3 tones occurring twice), 8 contour changes, and a pitch range of 11 semitones. None of the melodies implied a traditional Western tonal key as determined by the Krumhansl-Schmuckler key-finding algorithm (see Krumhansl, 1990), i.e., no significant correlation with any of the 24 tonal hierarchies. Melodies were isometric and isochronous, with an interonset interval of 250 ms. The surface-level groups (as defined by the I-R model) were consistently three tones in length, leading to an implied triple meter.

Procedure. On each trial, participants were presented with one of the eight test melodies. Following presentation of the melody, the participant was provided with five icons corresponding to a temporal sequence. The first icon was identified with the letter "T" and was always associated with the first three notes of the melody. The remaining icons (identified with letters "N," "P," "V," and "R") were associated with a unique three-note sequence drawn from the melody. Icons could be arranged in any sequence and played back at will. The initial presentation order of icons was randomized with the provision that at least one icon move was required before reconstructing the melody. Once participants believed they had correctly rearranged the icons, they were required to transcribe the letter tags onto an answer sheet. Trial orders were independently randomized for each participant.

RESULTS AND DISCUSSION

Reconstruction accuracy was at ceiling (95.8%) and thus not interpretable, but variability was present in: (1) number of icon moves; (2) number of replays; and (3) response latency (i.e., the time from the end of the initial presentation to the final transcription). A mixed analysis of variance was conducted on each of these three measures with emergent-level structure and surface-level structure as the within-subjects variables, and musicianship as the between subjects variable. An alpha-level of .05 was used for all statistical tests.

For each measure, the main effects of emergent-level structure and surface-level structure were significant. There were no significant main effects or interactions involving musicianship. Figure 4 displays the mean number of icon moves, replays, and response times, collapsed across musicianship.

Number of icon moves. Melodies with simple emergent-level structure received fewer icon moves than melodies with complex emergent-level structure, $F(1, 24) = 16.66$, $p < .001$. Melodies with simple surface-level structure received fewer icon moves than melodies with complex surface-level structure, $F(1, 24) = 9.89$, $p < .01$.

Replay. Melodies with simple emergent-level structure were replayed fewer times than melodies with complex emergent-level structure, $F(1, 24) = 7.21$, $p < .05$. Melodies with simple surface-level structure were replayed fewer times than melodies with complex surface-level structure, $F(1, 24) = 4.82$, $p < .05$.

Response time. Melodies with simple emergent-level structure led to shorter response times than melodies with complex emergent-level structure, $F(1, 24) = 5.64$, $p < .05$. Melodies with simple surface-level structure led to shorter response times than melodies with complex surface-level structure, $F(1, 24) = 4.39$, $p < .05$.

The results of these analyses support our prediction that melodic processing is facilitated by complexity at both emergent-level and surface-level of structure as defined by the I-R model. Although the predicted interaction between surface and emergent-level structure was not found, a consistent trend may be observed in Figure 4, wherein the advantage of simple emergent-level structure appears to be more present in melodies with simple surface-level structure.

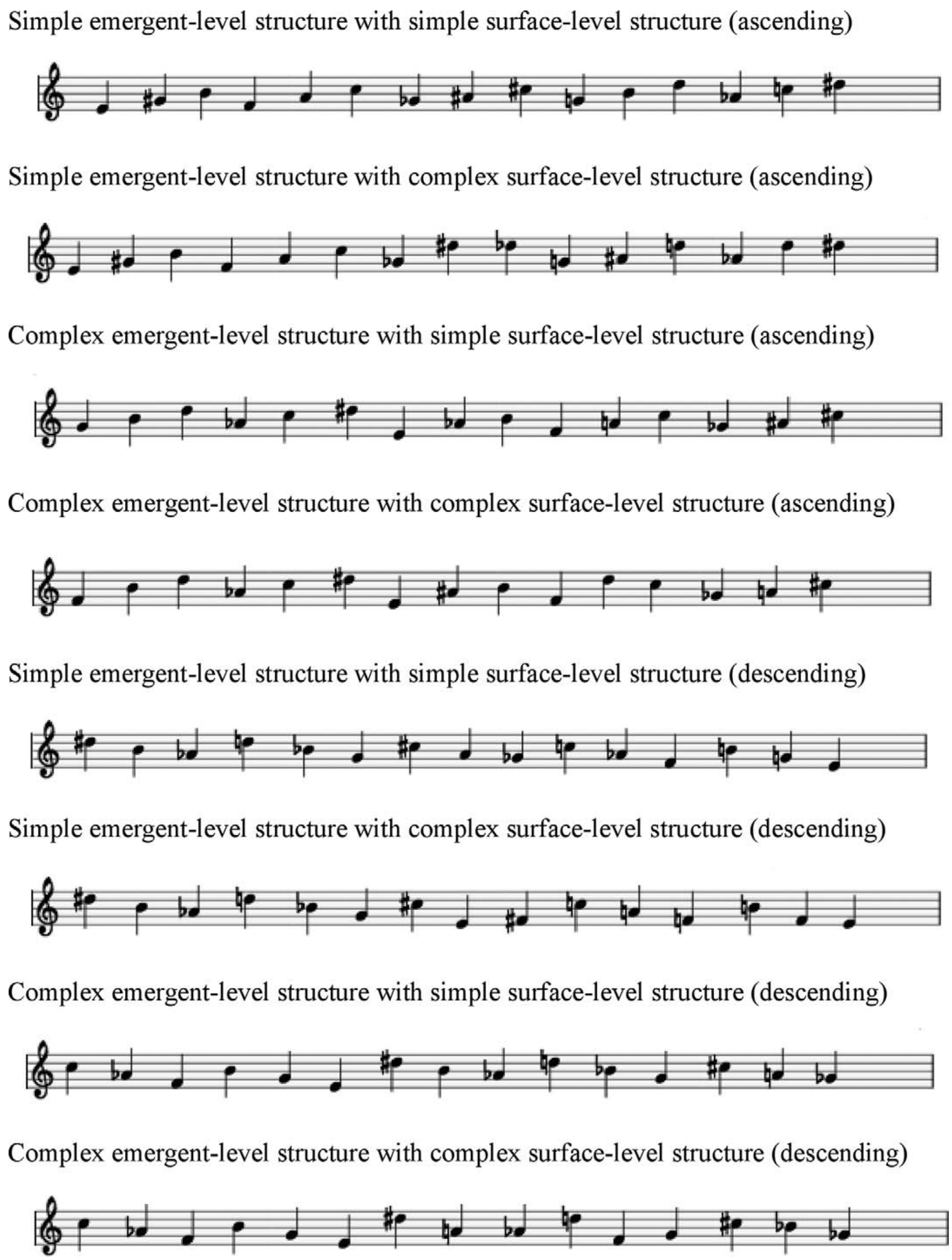

FIGURE 3. Musical notation for ascending (original) and descending (inverted) melodies used in Experiments 1 and 2.

Experiment 2

Much like the shape of a visually presented object, melodies are usually perceived in a "Gestalt" manner (von Ehrenfels, 1937). It is possible that the serial reconstruction task employed in Experiment 1 somehow altered this normative mode of listening. In Experiment 2, we adopted a cohesion-rating task where cohesion was defined as "the extent to which the tones of a melody sound as though they create an organized

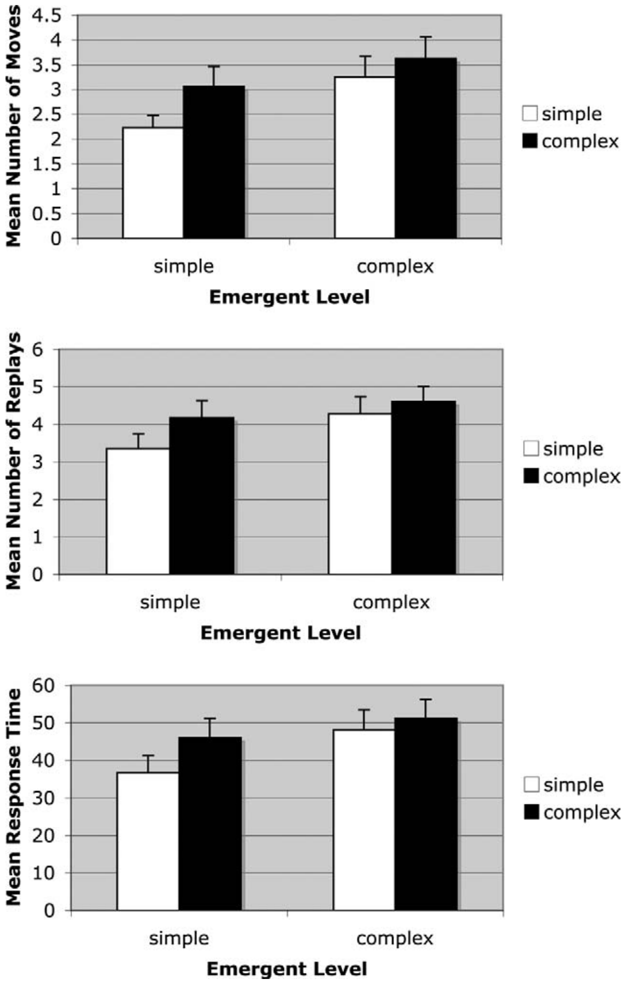

FIGURE 4. Mean number of icon moves, replays, and response times in Experiment 1 as a function of emergent-level complexity (simple vs. complex) and surface-level complexity (simple vs. complex).

whole." The operative assumption here is that judgments of cohesion will be influenced by ease of processing. Melodies were identical to those used in Experiment 1.

METHOD

Participants. Thirty undergraduate students were recruited to participate from the Queen's University Psychology Participant Pool. Demographic information is provided in Table 1. All participants were given course credit for their participation.

Apparatus. The cohesion-rating task was conducted in a sound-attenuated chamber with groups of 1 to 3 participants. Melodies were generated and presented over a Roland FP1 digital piano, set to "Piano 1," under the control of a Power Mac Computer running MusicShop software. Response sheets were used to record cohesion ratings.

Stimuli. Test melodies were identical to the eight melodies used in Experiment 1.

Procedure. Eighteen randomized orders of melody presentation were constructed, one for each of 18 test groups. Participants in each test group were asked to rate the cohesion of the eight melodies. Each melody was presented twice in succession with a 2-s pause between presentations. The cohesion of each melody was rated using a 7-point scale (1 = *not cohesive*, 7 = *very cohesive*). Participants were encouraged to use the full range of the scale. To familiarize participants with the nature of the melodies and the rating scale, participants were given two practice trials. The melodies in these practice trials were randomly selected from the set of test melodies.

RESULTS AND DISCUSSION

A mixed analysis of variance was conducted with emergent-level structure (simple vs. complex) and surface-level structure (simple vs. complex) as the within-subjects factors and musicianship as the between-subjects factor. Consistent with Experiment 1, main effects of emergent-level and surface-level structure were both significant. The interaction between emergent-level and surface-level structure was also significant.

Figure 5 displays cohesion ratings collapsed across musicianship. Melodies with simple emergent-level structure yielded higher cohesion ratings than melodies with complex emergent-level structure, $F(1, 28) = 77.96$, $p < .0001$. Melodies with simple surface-level structure yielded higher cohesion ratings than melodies with complex surface-level structure, $F(1, 28) = 91.37$, $p < .0001$. The advantage of simple emergent-level structure was amplified for melodies with simple

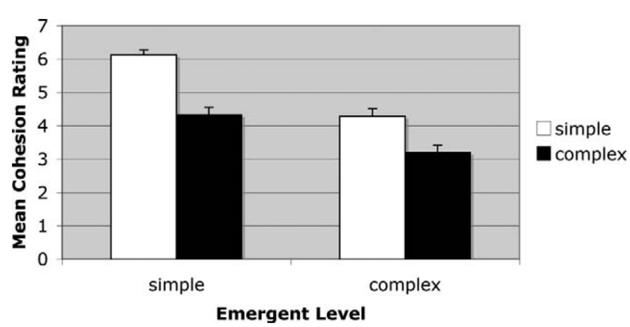

FIGURE 5. Mean cohesion ratings in Experiment 2 as a function of emergent-level complexity (simple vs. complex) and surface-level complexity (simple vs. complex).

surface-level structure, $F(1, 28) = 10.31$, $p < .003$. Although this interaction did not reach significance in the serial reconstruction data (Experiment 1), similar trends were apparent.

Although none of the melodies tested in the first two experiments were associated with a major or minor key, they all began with a prototypical diatonic pattern (major or minor triads) that was immediately repeated in transposition (A^0, A^0). The familiarity of these patterns may have helped to instantiate an implied triple meter, supporting perception of the emergent-level structure that was otherwise defined by systematic placement of expectancy denials. In addition, this implied meter may have been reinforced due to the lack of variability in the length of surface-level groups. Experiment 3 was conducted to explicitly control for these factors that may have contributed to the emergent-level findings.

Experiment 3

The results of Experiments 1 and 2 suggest that relations between emergent-level tones can influence the manner in which listeners perceive and remember melodies. However, the melodies used in these experiments contained structural cues beyond regular expectancy denials that may have reinforced emergent-level structure. Experiment 3 was conducted to determine whether sensitivity to emergent-level structure would persist when groups were not reinforced by these other cues.

METHOD

Participants. Thirty-two undergraduate students were recruited to participate from the Queen's University Psychology Participant Pool. Demographic information is provided in Table 1. Participants were given course credit for their participation.

Apparatus. The experiment was conducted in a sound-attenuated chamber at Queen's University. A Power Mac computer running Experiment Creator Software was used to present melodies and collect responses. Melodies were realized as MIDI performances using a piano patch, with sound output over Sennheiser HD280 headphones.

Stimuli. Sixteen melodies were composed for this experiment and presented in original and inverted form to create 32 test melodies (see Figure 6). None of the melodies contained prototypic triadic patterns. Melodies varied in emergent-level complexity (simple vs. complex), the number of tones in each surface level group (3 or 4 tones), and surface-level complexity (4 levels). As in prior experiments, emergent-level structure involved a sequence of small intervals forming a *process* (simple) or a combination of structures (complex). The four levels of surface-level complexity were created by manipulating the degree of redundancy between surface-level groups. At the simplest level (1), a single surface-level group was repeated (in transposition) throughout the melody. Higher levels of complexity progressively reduced the extent of repetition. The highest level of complexity (4) contained no repetition.

Eight dummy melodies were interspersed among the twenty-four test melodies. Dummy melodies were composed with surface-level groups that were five tones in length. The purpose of the dummy melodies was to reduce the likelihood that listeners would carry over expectations about group length from earlier trials. The resultant 32 melodies (24 test melodies and 8 dummy melodies) possessed the same number of surface-level groups (5), but varied in number of tones because of differences in group length (triple, quadruple, quintuple).

All melodies had a pitch range of 11 semitones with a frequency distribution of pitches that did not clearly imply a traditional Western tonal key as determined by the Krumhansl-Schmuckler key-finding algorithm (as described in Krumhansl, 1990). Melodies were isometric and isochronous, with an interonset interval of 250 ms.

Procedure. The procedure was identical to that described in Experiment 2 except that participants were run individually.

RESULTS AND DISCUSSION

A mixed analysis of variance was conducted with emergent-level structure (simple vs. complex), surface-level structure (4 levels of complexity), and group length (3 or 4 tones) as the within-subjects variables and musicianship as the between subjects variable. Figure 7 displays cohesion ratings collapsed across musicianship.

Melodies with simple emergent-level structure yielded higher cohesion ratings than melodies with complex emergent-level structure, $F(1, 29) = 11.53$, $p < .01$. Lower levels of surface-level complexity also led to higher cohesion ratings, $F(1, 29) = 18.23$, $p < .0001$. This finding is compatible with the main effect of surface-level structure revealed in Experiments 1 and 2.

We predicted that lower levels of surface-level complexity would emphasize emergent-level structure and amplify the advantage of simple emergent-level structure. Although the interaction between emergent-level structure and surface-level structure was not significant, $F(3, 87) = 1.44$, $p = .24$, the advantage of simple emergent-level structure was significant at all levels of

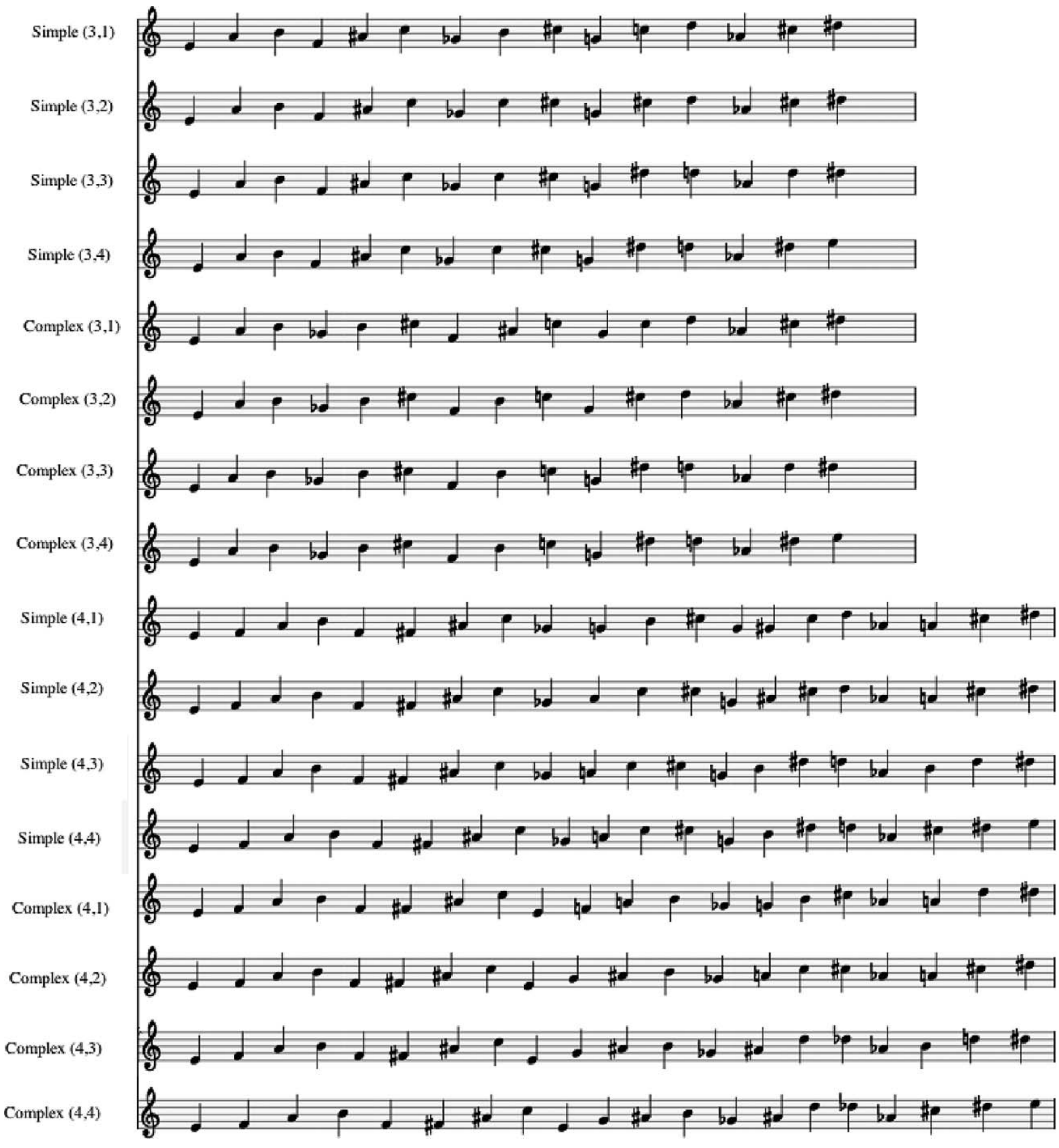

FIGURE 6. Musical notation for test melodies used in Experiment 3 (inverted melodies not shown). Each melody is labeled as having simple or complex emergent-level structure. The first number in brackets represents the number of notes in surface-level groups (3 or 4), while the second number represents the extent of surface-level complexity (1 = low and 4 = high).

surface-level complexity (all p values < .05), *except* for the highest level, $F(1, 29) < 1$. This finding suggests that it is not necessary to have strict repetition of the same surface-level group (in transposition) in order to observe effects of emergent-level structure.

There was an effect of group length, $F(1, 29) = 8.46$, $p < .01$. Melodies with groups that were four tones in length were perceived as more cohesive than melodies with groups that were three tones in length. One possibility is that this finding reflects a cultural bias in favor

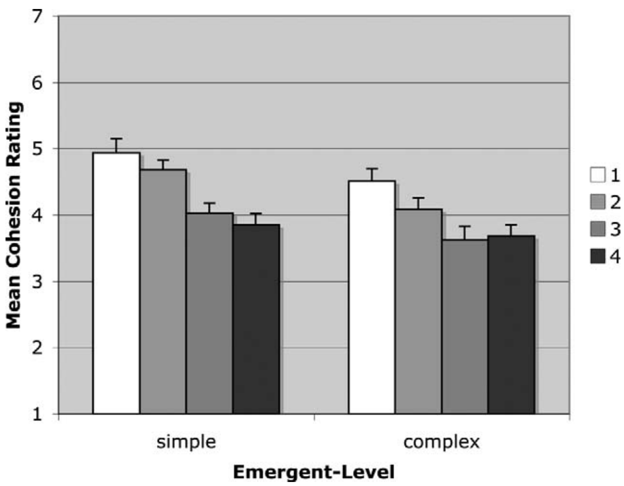

FIGURE 7. Mean cohesion ratings in Experiment 3 as a function of emergent-level complexity (simple vs. complex) and surface-level complexity (1 = low and 4 = high).

of duple time (see e.g., Smith & Cuddy, 1989; Trainor & Corrigall, 2010). Another possibility is that listeners' ratings were partly influenced by the absolute length of surface-level groups, with longer groupings deemed to be more cohesive.

Although melodies in this experiment were composed without the use of overly familiar surface-level groups, the first surface-level group was always repeated in transposition. The I-R model suggests that this initial repetition (A^0, A^0) should facilitate the perception of emergent-level structure (Narmour, 2000). Experiment 4 was designed to assess whether sensitivity to complexity of emergent-level structure could be observed using melodies that do not repeat the initial surface-level group.

Experiment 4

Experiments 1-3 revealed that ease of melodic processing depends on surface and emergent-level grouping. Surface-level groups were defined through the use of regularly occurring expectancy denials and by an immediate repetition of the initial surface-level group. The question addressed in Experiment 4 was whether this initial repetition is essential for the observed effect of emergent-level structure.

METHOD

Participants. Thirty undergraduate students were recruited to participate from the University of Toronto Community. Demographic information for each participant group (musician/nonmusician) is provided in Table 1. Participants recruited through the Introductory Psychology Participant Pool were given course credit for their participation. These participants included a mix of musicians and nonmusicians. Additional participants for the musician group were recruited using posters displayed around campus. Participants recruited using posters were reimbursed with nominal payment.

Apparatus. The experiment was conducted in a sound-attenuated chamber at the University of Toronto, Mississauga. The equipment used to present stimuli and collect data was identical to that described in Experiment 3.

Stimuli. Twelve melodies were composed and presented in both original and inverted forms to create 24 test melodies. Melodies varied in the number of tones in each surface-level group (3 or 4 tones), emergent-level structure (simple vs. complex), and surface-level complexity (3 levels). As may be seen in Figure 8, increasing levels of surface-level complexity were associated with lower levels of redundancy between surface-level groups, but in no case did this redundancy involve an immediate repetition of a melodic group (A^0, A^0). At the simplest level (1), two surface-level groups that formed a near repetition were alternated (A^a, A^b, A^a, A^b, A^a). At the most complex level (4), there was almost no repetition present across surface-level groups (A^a, B, A^b, C, D). As in Experiment 3, eight dummy melodies with surface-level groups of 5 tones were interspersed among the test melodies in order to minimize any carry-over effect of group length. All melodies possessed the same number of surface-level groups (5), but melodies varied in length from 15-25 tones because of the variable length of groups. All other aspects of the melodies were consistent with test melodies used in Experiment 3.

Procedure. The procedure was identical to that described in Experiment 3.

RESULTS AND DISCUSSION

A mixed analysis of variance was conducted with emergent-level structure (simple vs. complex), surface-level structure (3 levels of complexity), and group length (3 or 4 tones) as the within-subjects variables and musicianship as the between subjects variable. The main effect of musicianship and its interactions were not significant.

Figure 9 displays mean cohesion ratings across levels of surface-level and emergent-level complexity. The main effect of emergent-level structure did not reach significance, $F(1, 28) = 1.87$, $p = .18$. Thus, eliminating

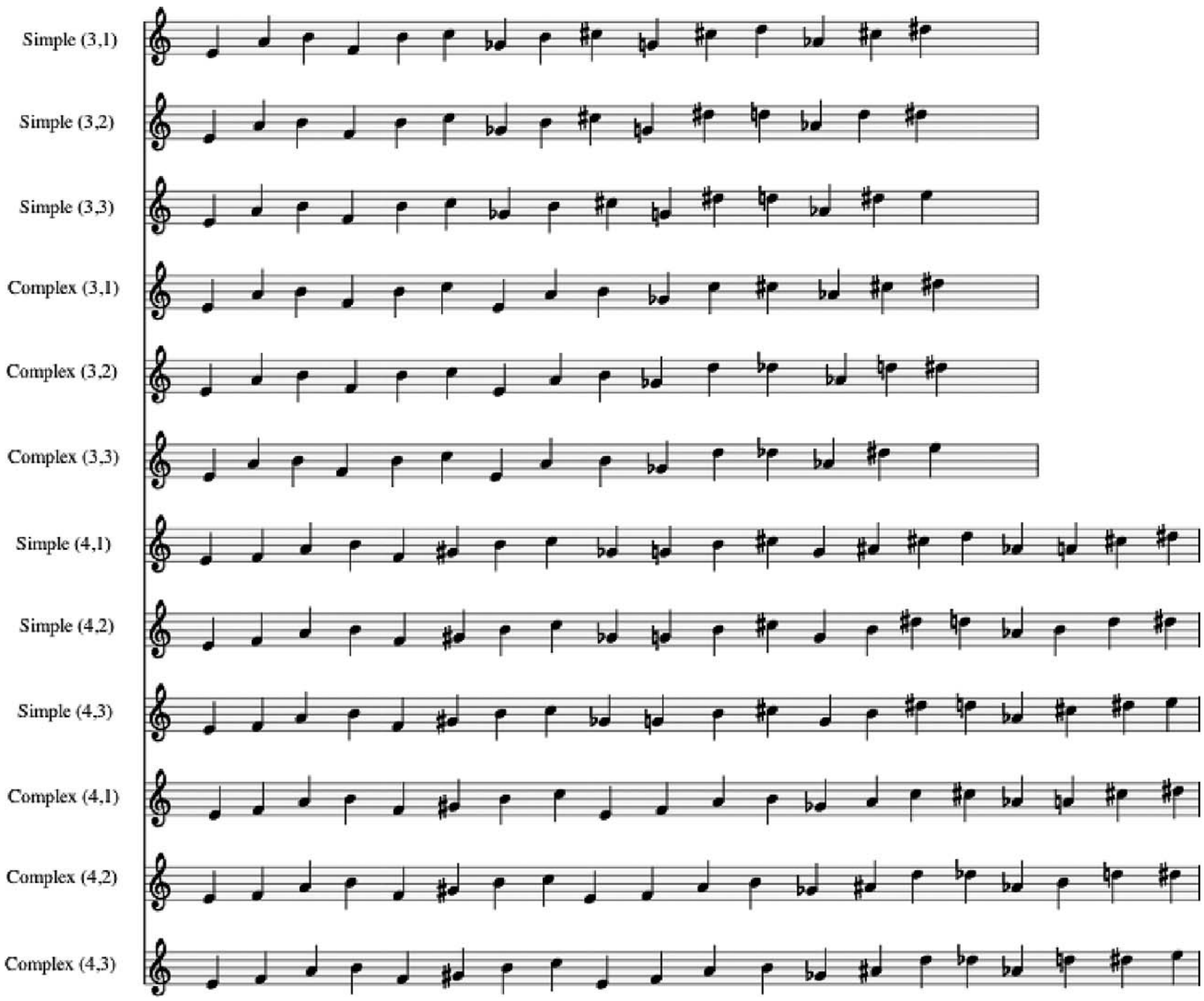

FIGURE 8. Musical notation for test melodies used in Experiment 4 (inverted melodies not shown). Each melody is labeled as having simple or complex emergent structure. The first numbers in brackets represents the number of notes in surface-level groups (3 or 4), while the second number represents the extent of surface-level complexity (1 = low and 3 = high).

repetition of the first surface-level group reduced the likelihood that listeners would perceive the emergent-level structure. However, the interaction between emergent-level structure and surface-level structure was significant, $F(2, 56) = 4.01$, $p < .05$. Orthogonal contrasts revealed that while there was no effect of emergent-level structure in melodies with high (3) and intermediate (2) surface-level complexity, $F(1, 28) < 1$, the emergent-level effect was significant in melodies with low (1) surface-level complexity, $F(1, 28) = 5.75$, $p < .05$. For melodies with low surface-level complexity, melodies with simple emergent-level structure were judged as more cohesive than melodies with complex emergent-level structure. Thus it seems that at least for nontonal melodies, some surface repetition may be necessary to perceive emergent-level effects.

The effects of surface-level complexity and group length were all consistent with those observed in Experiment 3. Ratings of cohesion were higher for melodies with simple surface-level structure, $F(2, 56) = 41.01$, $p < .0001$. This finding was predicted and is attributable to the reduction in complexity that is conferred by increasing pattern repetition. Ratings were higher for melodies with surface-level groups that were four tones in length, $F(1, 28) = 16.60$, $p < .0001$. As suggested in Experiment 3, possible explanations for this finding include a cultural bias in favor of duple time or an effect of absolute length of surface-level groups. Although this issue is beyond the

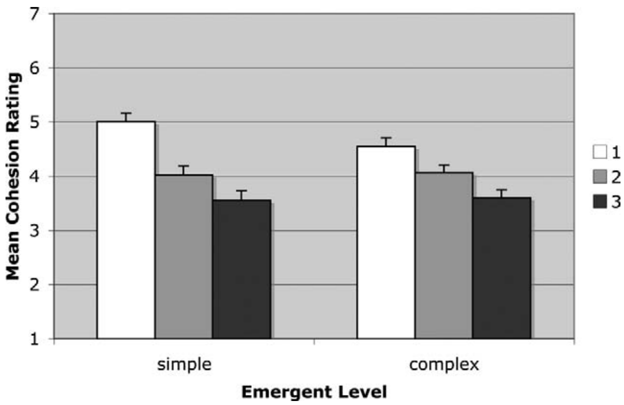

FIGURE 9. Mean cohesion ratings in Experiment 4 as a function of emergent-level complexity (simple vs. complex) and surface-level complexity (1 = low and 3 = high).

scope of this study, future research might contrast these two interpretations by including melodies with even longer groupings that are not in duple time (e.g., five-note groups).

General Discussion

Serial-reconstruction and cohesion-rating tasks were administered to assess processing difficulty in hierarchically structured melodies as defined by the I-R model. There were three main findings. First, melodies with simple emergent-level structure were easier to process than melodies with complex emergent-level structure. Second, melodies with simple surface-level structure were easier to process than melodies with complex surface-level structure. Third, sensitivity to emergent-level structure generally increased with increasing simplicity at the surface level. None of the experiments yielded main effects or interactions involving musicianship.

The processing advantage for melodies with simple emergent-level structure has been characterized using music-theoretic descriptions of emergent-level structure derived from the Implication-Realization model. In such melodies, a single emergent-level structure connected together non-adjacent tones that were the first tones of surface-level groups. In the terms of the I-R model, the emergent-level structure formed a *process*, involving sequential pitch intervals composed of non-adjacent tones moving in a common direction. A process is considered a highly expected structure by Narmour (1989, 1990, 1999), as well as other theorists (Huron, 2006; Larson, 2012; Margulis, 2005; Meyer, 1956, 1973).

On the other hand, the processing advantage of melodies with simple emergent-level structure could also be argued from an information-theoretic perspective. Melodies that can be described with relatively short codes are easier to process than melodies with longer codes. However, while it is true that an emergent-level process can be captured by a short code, coding systems do not provide a mechanism for the establishment of emergent-level structure in the absence of exact repetition at the surface level.

Although an expectancy-based explanation for the emergent-level findings appears likely, there were no direct tests of expectancy conducted in this study. Both experimental tasks required that listeners make retrospective judgments (serial reconstruction and cohesion). Hence, it remains remotely possible that some account of simplicity that does not depend upon expectation per se was responsible for the pattern of judgments observed. It would be valuable for future studies of hierarchical melodic structure to combine prospective tasks that tap expectancy alongside of retrospective tasks. It would also be useful to incorporate melodies that fulfill emergent-level expectancy in a manner that is distinct from process (e.g., a *Narmourian* reversal).

Sensitivity to emergent-level structure became more apparent when multiple structural cues were available. The presence of familiar patterns (Experiment 1, 2), an immediate repetition (Experiment 3), and/or minimal variation in surface-level groups (Experiment 4; Surface Level Complexity = 1), appear to have been instrumental in yielding effects of emergent-level structure. We do not presume that this is an exhaustive list of criteria but it seems that a listener's attention will tend to remain at the melodic surface in the absence of substantive evidence reinforcing a possible emergent-level structure.

Other studies have also found evidence for sensitivity to emergent-level structure in melody. Memory for short and simple melodic structures appears to preserve emergent-level structure (Bigand, 1990, Exp. 3; Deutsch & Feroe, 1981; Sloboda & Parker, 1985). Listeners prefer correct over incorrect melodic reductions (Dibben, 1994; Serafine et al., 1989), and are able to perceive tonal tension in non-adjacent dependencies (Lerdahl & Krumhansl, 2007; see also Cuddy & Smith, 2000; Smith & Cuddy, 2003).

Statistical learning studies have also revealed sensitivity to non-adjacent dependencies. By manipulating pitch proximity between the odd and even tones in a melody, Creel, Newport, and Aslin (2004) were able to affect grouping such that relationships between non-adjacent tones were readily learned. On the basis of this finding and those of the present study, it seems reasonable to predict that statistical learning of non-adjacent dependencies will be harder in melodies that deny emergent-level expectancy. In this way, the bottom-up principles may form a sort of bottleneck for the processing of melodic structures (Rohrmeier & Cross, 2013).

Other studies have cast doubt on listeners' sensitivity to emergent-level structure. Cook (1987) investigated listeners sensitivity to tonal closure. Listeners were asked to indicate their preference between two versions of the same piano excerpt, only one of which involved a return to the initial key. Results showed little effect of tonal closure on judgments. A similar result was obtained by Marvin and Brinkman (1999) when they explicitly asked participants about whether excerpts ended in the key with which they started.

In addition, Bigand and Parncutt (1999) found no evidence for an influence of hierarchical structure in chord sequences on judgments of tension. This finding stands in contrast with tension data reported by Lerdahl and Krumhansl (2007). One explanation for this discrepancy is that tension was defined differently across the two studies. Bigand and Parncutt (1999) defined tension as a feeling of non-closure in the sense that "there must be a continuation of the sequence" (p. 242). Lerdahl and Krumhansl (2007) did not emphasize this nonclosural aspect in their definition of tension. It seems that there are many instances in which these two definitions may lead to unique modes of listening and hence contradictory outcomes – e.g., the initial tonic chord of a harmonic sequence where there is no implication of closure.

The picture that is developing from various studies of emergent-level structure is that listeners can be rather sensitive to theoretical descriptions of structure beyond the surface, but that the extent of this sensitivity depends greatly on the cues supporting the hierarchy and on the listening mode. The current investigation adds to existing research on melodic perception by demonstrating that emergent-level structure contributes to processing difficulty and that emergent-level structure may be instantiated in the absence of exact repetition at the surface.

Author Note

This research was supported by independent discovery grants awarded to each author from the Natural Sciences and Engineering Research Council of Canada. We thank Eugene Narmour, Daniel Levitin, and Aniruddh Patel for helpful suggestions.

Correspondence concerning this article should be addressed to Frank Russo, Department of Psychology, Ryerson University, Toronto, Ontario, M5B 2K3, Canada. E-mail: russo@ryerson.ca

References

Bigand, E. (1990). Abstraction of two forms of underlying structure in a tonal melody. *Psychology of Music, 18*, 45-59.

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research, 62*, 237-254.

Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology, 18*, 389-431.

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception, 5*, 197-206.

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1119-1130.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 869-883.

Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception and Psychophysics, 57*, 451-462.

Cuddy, L. L., & Smith, N. A. (2000). Perception of tonal pitch space and tonal tension. In D. Greer (Ed.), *Musicology and sister disciplines: Past, present, future* (pp. 47-59). Oxford: Oxford University Press.

Deliège, I., & Mélen, M. (1997). Cue abstraction in the representation of musical form. In I. Deliège & J. A. Sloboda (Eds.), *Perception and cognition of music* (pp. 387-412). London: Lawrence Erlbaum Associates.

Deliège, I., & Mélen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real time listening to a piece of music. *Music Perception, 14*, 117-160.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception and Psychophysics, 28*, 381-389.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88*, 503-522.

Dibben, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *Music Perception, 12*, 1-25.

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.

Forte, M. R. A. (1977). *The structure of atonal music*. New Haven, CT: Yale University Press.

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience, 16*, 1010-1021.

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics*, *41*(6), 621-634.

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

Jones, M. R., & Boltz, M. (1989). Dynamic attention and responses to time. *Psychological Review, 96*, 459-491.

Jones, M. R., & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *51*(4), 271-291.

Jones, M. R., & Ralston, J. T. (1991). Some influences of accent structure on melody recognition. *Memory and Cognition*, *19*(1), 8-20.

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

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

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119-159.

Larson, S. (2012). *Musical forces: Motion, metaphor, and meaning in music*. Bloomington, IN: Indiana University Press.

Leeuwenberg, E. L. (1969). Quantitative specification of information in sequential patterns. *Psychological Review, 76*, 216-220.

Lerdahl, F. (1988). Tonal pitch space. *Music Perception, 5*, 315-350.

Lerdahl, F. (1989). Atonal prolongational structure. *Contemporary Music Review, 4*, 65-87.

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

Lerdahl, F., & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception, 24*, 329-366.

Loui, P. (2012). Learning and liking of melody and harmony: Further studies in artificial grammar learning. *Topics in Cognitive Science, 4*, 1-14.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception, 21*, 663-714.

Margulis, E. H. (2013). *On repeat: How music plays the mind.* New York: Oxford University Press.

Marvin, E. W., & Brinkman, A. (1999). The effect of modulation and formal manipulation on perception of tonic closure by expert listeners. *Music Perception, 16*, 389-407.

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

Meyer, L. B. (1973). *Explaining music*. Berkeley, CA: University of California Press.

Narmour, E. (1983). Some major theoretical problems concerning the concept of hierarchy in the analysis of tonal music. *Music Perception*, *1*, 129-199.

Narmour, E. (1989). The "genetic code" of melody: Cognitive structures generated by the implication-realization model. *Contemporary Music Review, 4,* 45-63.

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

Narmour, E. (1991). The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception, 9*, 1-26.

Narmour, E. (1992). *The analysis and cognition of melodic complexity: The implication-realization model.* Chicago, IL: University of Chicago Press.

Narmour, E. (1999). Hierarchical expectation and musical style. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 441-472). San Diego, CA: Academic Press.

Narmour, E. (2000). Music expectation by cognitive rule-mapping. *Music Perception, 17*, 329-398.

Pearce, M. T., Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception, 23*, 377-405.

Pearce, M. T., Wiggins, G. A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science, 4*, 625-652.

Rohrmeier, M., & Cross, I. (2013). Artificial grammar learning of melody is constrained by melodic inconsistency: Narmour's principles affect melodic learning. *PloS One, 8*, e66174.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review*, *77*(6), 481.

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition, 58*, 75-125.

Schellenberg, E. G. (1997). Simplifying the implication-realization model of melodic expectancy. *Music Perception, 14*, 295-318.

Schenker, H. (1969). *Five graphic musical analyses*. New York: Dover Publications.

Schenker, H. (1979). *Free composition* [Der freie Satz] (E. Oster, Trans. & Ed.). New York: Pendragon Press.

Serafine, M. L., Glassman, N., Overbeeke, C. (1989). The cognitive reality of hierarchic structure in music. *Music Perception, 6*, 397-430.

Simon, H. A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review, 79*, 369-382.

Sloboda, J. A., & Parker, D. H. H. (1985). Immediate recall of melodies. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 143-167). London: Academic Press.

Smith, K. C., & Cuddy, L. L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 457-471.

Smith, N. A., & Cuddy, L. L. (2003). Perceptions of musical dimensions in Beethoven's Waldstein sonata: An application of tonal pitch space theory. *Musicae Scientiae*, *7*, 7-34.

Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research, 61*, 157-174.

Trainor, L. J., & Corrigall, K. A. (2010). Music acquisition and effects of musical experience. In M. R. Jones, R. R. Fay, and A. Popper (Eds.), *Music perception* (pp. 89-127). New York: Springer.

von Ehrenfels, C. (1937). On Gestalt-qualities. *Psychological Review, 44*, 521-524.

von Hippel, P. (2002). Melodic-expectation rules as learned heuristics. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, J. Renwich (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition*. Sydney, Australia: ICMPC.

▼ Basic Description

Title	Music Perception
ISSN	0730-7829
Publisher	University of California Press * Journals Division
Country	United States
Status	Active
Start Year	1983
Frequency	5 times a year
Volume Ends	Summer
Language of Text	Text in: English
Refereed	Yes
Abstracted / Indexed	Yes
Serial Type	Journal
Content Type	Academic / Scholarly
Format	Print
Website	http://www.ucpressjournals.com/journal.asp?j=mp
Email	mpercep@queensu.ca
Description	Focuses on scientific and musical approaches to the study of musical phenomena.

▶ Subject Classifications

▶ Additional Title Details

▶ Publisher & Ordering Details

▶ Price Data

▶ Online Availability

▶ Abstracting & Indexing

▶ Other Availability

▶ Demographics

▶ Reviews

Related Titles

▶ **Alternative Media Edition** (2)