
Perceiving the Tonal Ending of Tune Excerpts: The Roles of Pre-existing Representation and Musical Expertise

SYLVIE HÉBERT, ISABELLE PERETZ, and
LISE GAGNON *Université de Montréal*

Abstract In this study, sensitivity to tonal relations was assessed by using real melodies instead of traditional scales or chords. Two groups of listeners — one trained, one untrained — rated the goodness of fit of each of the 12 tones of the chromatic scale as the final note of familiar and unfamiliar tune excerpts. The unfamiliar excerpts were the mirror forms in pitch and time of the familiar tunes. The results showed that musicians and nonmusicians exhibited responses that were governed by tonal relations with both familiar and unfamiliar tunes. These findings were corroborated by multiple-regression analysis, which revealed that the pattern of ratings reflected knowledge of the musical structure, beyond the contribution of surface features such as note frequency or pitch proximity between the two last tones.

When listening to Western music, whether classical or popular, even the most naive listener can get the feeling of “grasping” the music unfolding. He or she can whistle a tune or a musical passage after having heard it only a few times, and can quite accurately predict when the piece will end. In short, any person who has been exposed from an early age to Western music can understand it without having explicit knowledge of its structure. Various sources of knowledge can influence people’s interaction with the music. In the present study, we will be concerned with three types of such knowledge: general knowledge of tonal structure, specific knowledge of tunes, and expert knowledge of music.

Tonal knowledge may be regarded as one of the central mental systems upon which the understanding of Western music depends. In its broadest sense, tonality is applied to musical compositions in which all notes appear to be related to a single, central note, the tonic. Almost every style of music, from Gregorian chant to modern music, could thus be called tonal. Yet, in a more restrictive sense, tonality entails the specific notions of scale and key. Most Western music is written in one particular key, which is indicated in the musical score by the key signature. That is, within a composition, tones

constituting the key are heard more often than the remaining tones (Krumhansl, 1990, p. 67). Further, they are heard at strategic points in time (Boltz, 1989, 1991; Jones & Boltz, 1989).

Listeners who share a common musical idiom have internalized these regularities through exposure to the music. They have created mental schemas that enable them to encode and remember tones easily and reliably. For instance, listeners find tonal sequences much easier to remember than nontonal sequences (Francès, 1958), and tonal sequences easy to distinguish from nontonal ones (Dowling, 1978). Likewise, single tones are easier to remember when embedded in tonal sequences than when presented in isolation or when presented in random sequences (Dewar, Cuddy & Mewhort, 1977).

Music theorists (e.g., Meyer, 1956; Lerdahl, 1988) have suggested that the knowledge base required to capture tonal regularities is organized in a hierarchical fashion. Key is said to serve as a psychological framework within which some tones are perceived as more stable or significant than others, with the tonic acting as a central pole around which all other chromatic scale degrees are more or less distantly related. Stability refers to a sense of completeness or resolution once a tonal context has been established (Bharucha, 1984). The tones that are the least stable are the most remotely related to the tonic; these are the nonscale tones. The second level of the tonal hierarchy involves a preference, or greater sense of stability, for the tones that make up the tonic triad, that is the first degree of the scale (or tonic), the third degree (or mediant), and the fifth degree (or dominant), over the other nontriad scale tones. Within the triad, the tonic is perceived as the most stable tone, and the dominant is perceived as more stable than the mediant; this relation composes the third and highest level of the tonal hierarchy.

The tonal hierarchy has been supported empirically with the probe-tone paradigm. The technique, devised by Krumhansl and Shepard (1979), consists of presenting subjects with a context, such as scales (Speer & Meeks, 1985; Cuddy & Baderstcher, 1987; Janata & Reisberg, 1988), triads (Cuddy & Baderstcher, 1987; Cuddy, 1991), chords (Janata & Reisberg, 1988), or chord progressions (Krumhansl & Kessler, 1982; Cuddy & Thompson, 1992), which establishes a specific key. The context is followed by one of the tones of the chromatic scale and subjects are asked to rate whether the probe “fits with” or “is a good completion for” the prior context. The obtained response profiles parallel the relative stability of each tone within the key, as suggested by music theory.

These conclusions have been drawn mostly from the results obtained with musicians (Janata & Reisberg, 1988; Krumhansl & Keil, 1982; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979; Frankland & Cohen, 1990). A problem in validating the tonal hierarchy in musicians is that the results can be the reflection of explicit reconstruction of learned tonal rules rather than the product of the more implicit knowledge of tonal structure that has been

acquired through mere exposure to tonal music. Interestingly, the few studies that have considered nonmusicians (Krumhansl & Shepard, 1979; Cuddy & Badertscher, 1987; Frankland & Cohen, 1990; Janata & Reisberg, 1988; Cuddy & Thompson, 1992) or children with little musical training (Krumhansl & Keil, 1982; Speer & Meeks, 1985; Cuddy & Badertscher, 1987) have sometimes failed to obtain the musicians' response profiles.

Nonmusicians' responses often appear to be governed by nontonal cues like pitch proximity to the last tone of the context (Krumhansl & Shepard, 1979; Cuddy & Badertscher, 1987; Janata & Reisberg, 1988; Frankland & Cohen, 1990) or pitch direction of the final interval (Peretz, 1993). To account for this relative lack of sensitivity to tonal hierarchy in listeners with little or no musical training, Krumhansl (1990) argues that these subjects may adopt a logical, "nonmusical" attitude in laboratory tasks. In other words, nonmusicians possess tonal schemas, but the laboratory tasks are not tailored to tap this implicit knowledge. A more drastic account has, however, been offered. Serafine (1983) has interpreted nonmusicians' negative results as evidence that tonal knowledge is an artifact of music theory or of musicians' application of learned rules, thus having little or no processing validity.

We favour, however, Krumhansl's view that tonal knowledge is indeed the product of experience acquired early in ontogenetic development even without explicit tutoring (as has been argued in more detail in Peretz & Morais, 1989), and that artificial task demands in laboratory settings often do not allow this knowledge to come into play. One factor that might compensate for this lack of ecological validity is the use of highly familiar material instead of the usual novel and musically impoverished material.

That familiarity with the material can indeed improve nonmusicians' performance has been observed by Attneave and Olson (1971) for the transposition of short melodic sequences. In that study, subjects were given a short melodic pattern and were asked to match it in a different frequency region. Nonmusicians, who failed on the task when the sequences were novel, succeeded when a well-known, highly overlearned musical sequence — the NBC chimes — was employed. We thus hypothesized that using highly familiar material with the probe-tone technique would be a good way to promote the use of procedures that are routinely used while listening to music.

Although we expected nonmusicians to perceive tonal structure in well-known tunes, we could not make any specific prediction about how they would perform with real, but unfamiliar, melodies. So far, no study has used real melodies. In contrast, we did expect musicians to show sensitivity to tonal hierarchy, in both familiar and unfamiliar contexts. Musicians are trained to be efficient in a variety of contexts, by applying general principles which help them to focus on cues that are tonally most relevant.

To test these predictions, seven familiar tunes, from which seven unfamiliar ones were derived, served as musical contexts in separate conditions. Familiar

tunes were taken from the popular repertoire of French Quebec natives and had been rated as highly familiar (Peretz, Babaï, Lussier, Hébert, & Gagnon, 1995). Unfamiliar tunes were retrograde inversions of the familiar tunes, inverting both pitch and time so that the last tone becomes first and the first becomes the last. Although retrograde versions may not equate the original tunes on all relevant dimensions, they are maximally close to the familiar ones in interval size, tone repetition, rhythmic structure and complexity.

For testing purposes, each context, either familiar or unfamiliar, was followed by one of the twelve tones of the chromatic scale. This probe tone was to be judged as either providing a good ending or not providing a good ending. The choice of a dichotomous response — “yes” or “no” — instead of the rating scale used in most prior studies (Krumhansl & Shepard, 1979; Krumhansl & Keil, 1982; Krumhansl & Kessler, 1982; Speer & Meeks, 1985; Cuddy & Badertscher, 1987; but see Frankland & Cohen, 1990; Janata & Reisberg, 1988 for exceptions), was introduced to simplify the decision-making process. A previous study of ours (Peretz, 1993) had suggested that this mode of response was preferable for nonmusicians. Four groups of subjects, two highly trained and two untrained, performed the task in one of the two conditions defined by tune familiarity.

METHOD

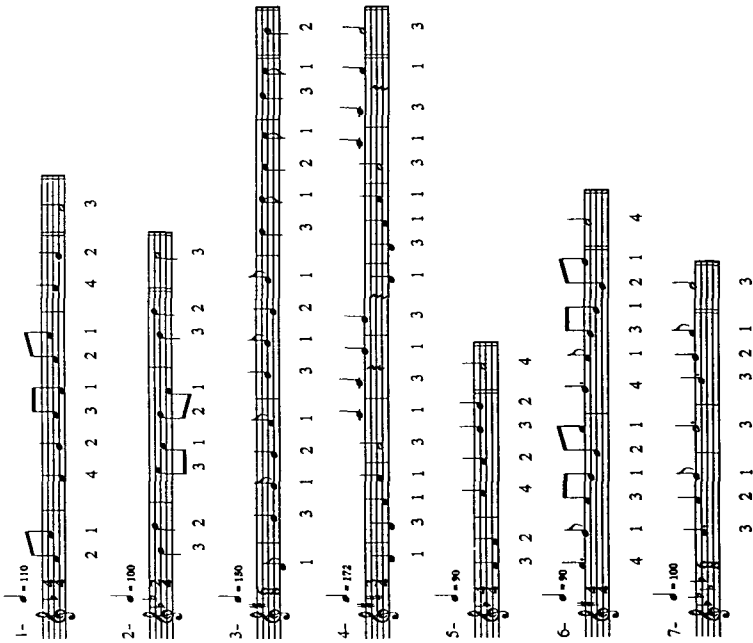
Subjects

Forty subjects (20 women and 20 men) having a mean age of 26.2 years (range: 18-35) were recruited mostly from the university community for the experiment. Twenty subjects were semi-professional musicians or graduate students in our Music Department (mean years of formal training in music theory = 5.9; mean years of practice of an instrument = 13.7). None reported having absolute pitch. The other twenty subjects were nonmusicians, having less than 3 years of musical training (mean years of formal training in music theory = 0; mean years of practice of an instrument = 1.2). Subjects were paid for their participation.

Stimuli and equipment

Seven familiar musical fragments were selected so that the final tone always occurred on a downbeat. The downbeat is any strong rhythmic point within a melody; for instance, in a waltz, the downbeats are the “ONES” of the ONE, two, three, ONE, two, three, etc. From a temporal point of view, cutting the fragment at this point optimized the sense of ending for the listener. Also, each final tone corresponded to a different degree of the major scale. That is, without regard to its specific key, each fragment ended on one of the seven possible tones which are part of the major scale. This procedure ensured that any memory effect for the real ending was distributed across the scale tones, because the real ending was not part of the context but was presented as one

Familiar contexts:



1- $\text{♩} = 110$
2 1 4 2 3 1 2 1 4 2 3

2- $\text{♩} = 100$
3 2 3 1 2 1 3 2 3

3- $\text{♩} = 150$
1 3 1 2 1 3 1 2 1 3 1 2

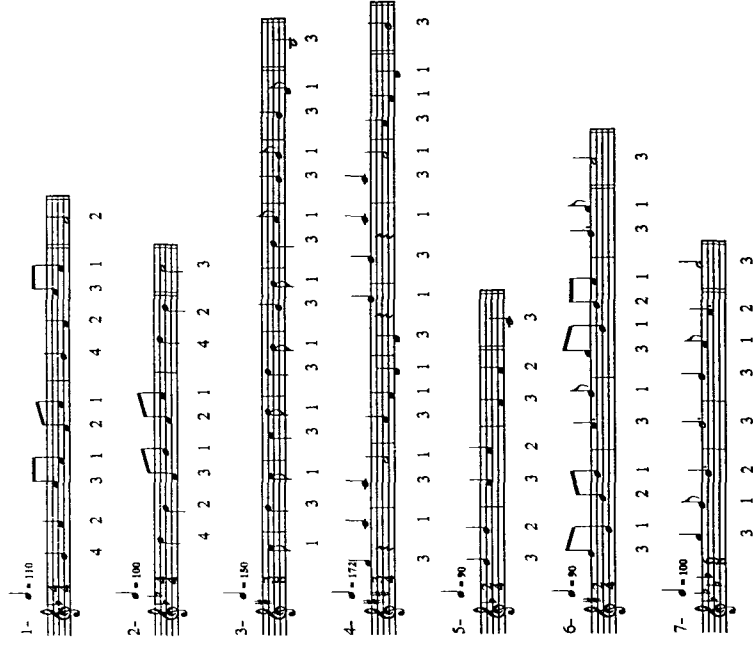
4- $\text{♩} = 172$
1 3 1 1 3 1 3 1 3 1 1 3 1 3

5- $\text{♩} = 90$
3 2 4 2 3 2 4

6- $\text{♩} = 90$
4 1 3 1 2 1 4 1 3 1 2 1 4

7- $\text{♩} = 100$
3 2 1 3 3 2 1 3

Unfamiliar contexts:



1- $\text{♩} = 110$
4 2 3 1 2 1 4 2 3 1 2

2- $\text{♩} = 100$
4 2 3 1 2 1 4 2 3

3- $\text{♩} = 150$
1 3 1 3 1 3 1 3 1 3 1 3 1 3

4- $\text{♩} = 172$
3 1 3 1 3 1 1 3 1 3 1 3 1 1 3

5- $\text{♩} = 90$
3 2 3 2 3 2 3

6- $\text{♩} = 90$
3 1 2 1 3 1 3 1 2 1 3 1 3

7- $\text{♩} = 100$
3 1 2 3 3 1 2 3

Figure 1. Familiar and unfamiliar tune excerpts used in the study. The familiar tunes display the “real ending”, which was presented as one of the possible probe tones.

of the possible probe tones. All the fragments were in major keys, and none of them had accidentals (i.e., tones that do not belong to the key). However, the fragments varied considerably in their number of tones (from 7 to 18 tones) and their tempo (from 90 to 172 beats/min). Therefore, their length also varied, from 4.8 s to 8.4 s (mean: 6 s). All fragments are presented in Figure 1. All were judged to be very familiar (with a mean of 4.8 on a 5-point scale, where 1 indicates very unfamiliar and 5 highly familiar, by 60 subjects for fragments 2, 4, 5, and 6, and a mean of 3 on a 4-point scale by 10 other subjects for fragments 1, 3, and 7). Moreover, when presented with the seven original excerpts before the beginning of the experiment, all the experimental subjects were able to name or continue to sing the excerpts correctly, thus confirming that they had access to the correct musical representations.

The unfamiliar musical fragments were the retrogrades in pitch and time of the familiar ones. To avoid shifting all tones of each bar and so as to respect the metrical structure of the familiar tunes, the unfamiliar fragment started with the penultimate tone of the familiar fragment. That is, the penultimate tone of the familiar fragment became the first of the unfamiliar fragment, the tone before the penultimate became the second, and so on up to the first tone of the familiar fragment. In other words, when one plays the unfamiliar tune backwards, one hears the integral familiar tune, minus its last tone. Three slight adjustments were made to the retrograde versions to soothe some awkward metrical consequences. Familiar fragment #4 in Figure 1 begins with an incomplete bar; keeping this in the retrograde would have unbalanced the metrical structure. Accordingly, this tone was dropped in the unfamiliar fragment. In the unfamiliar fragment #6, the duration of tones 5 and 6 were interchanged to make the fragment sound rhythmically smoother. The same procedure was applied in fragment #7 to tones 1 and 2 and to tones 5 and 6.

The seven unfamiliar sequences were played to each of the experimental subjects prior to the experiment. None of the subjects reported recognizing any of the fragments.

These two sets of tune fragments served as musical contexts in the familiar condition and the unfamiliar condition, respectively. Each trial consisted of the presentation of a randomly selected tune fragment, followed by a probe tone drawn randomly from the 12 chromatic tones in the corresponding octave. The pitch height of the probe tone was chosen to be close to the pitch height of the last tone. That is, the probe tone was within a distance of an octave from the last tone of the fragment. Thus, half of the probe tones were higher in pitch than the preceding tone, and half were lower. The pitch direction of the last interval was counterbalanced across the probe tones of each context. There were 84 different trials for each condition defined by tune familiarity. This series was repeated once, but in a different order, totalling 168 trials per condition.

The material was generated by a Yamaha TX-81Z synthesizer controlled by

an IBM-compatible computer. The analog output was pre-set to a timbre imitating the sound of a pan flute, and was recorded with a B77 Revox tape recorder. The tapes were played back to the subject via Uher W 766 A headphones at a comfortable listening level.

Procedure

Half of the musicians and half of the nonmusicians were randomly assigned to each of the two familiarity conditions. Each subject was tested individually in a session lasting about 60 min, including a 15-min pause after the first half of the experiment. The experimental session was preceded by seven examples, comprising each of the different fragments with different types of probe tones. Subjects were instructed to judge whether the last tone of the fragment was or was not an acceptable ending in the context of the melody. They responded by moving a bi-directional response stick either toward or away from themselves if their response was "yes" or "no", respectively. No feedback was provided to the subjects.

RESULTS

The data were subjected to two different types of statistical analyses: an overall analysis of variance to assess the role of musical expertise and of tune familiarity, and several regression analyses to assess the contribution of various structural cues to the responses.

Variance analyses

The scores for each subject were entered as follows: All probe tones corresponding to the tonic were coded as Probe-Tone Position 0; the other probe tones, which represent successive chromatic steps from the tonic, were coded as Probe-Tone Positions 1 to 11. For each one of these 12 Probe-Tone Positions, the proportion of "yes" responses, collapsed across the two blocks of trials and the seven fragments, was converted into a percentage and taken as the dependent measure.

An analysis of variance was carried out on these data with Musical Expertise and Tune Familiarity as the between-subjects variables and Probe-Tone Position as the within-subjects variable. This analysis revealed a main effect of Probe-Tone Position, $F(11,396) = 129.83$, $MS_e = 4.09$, $p < .001$, and also an interaction between Tune Familiarity and Probe-Tone Position, $F(11,396) = 4.05$, $MS_e = 4.09$, $p < .001$. There was no effect of Musical Expertise or Tune Familiarity, both $F_s < 1$. No interaction was found between Musical Expertise and Tune Familiarity, $F(1,36) = 1.32$, $MS_e = 25.41$. These results point to a high degree of agreement among the ratings of the musicians and the nonmusicians. They also indicate that the responses fluctuate as a function of probe-tone position, both within and between the two conditions. Further analyses were then conducted to determine the

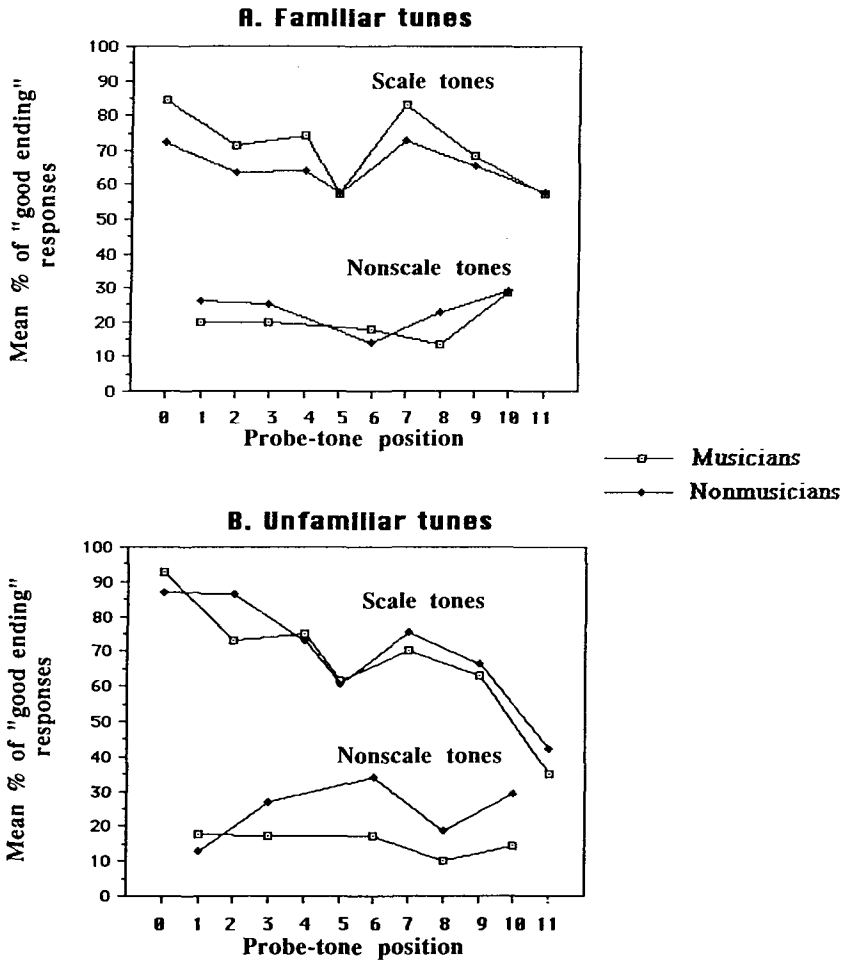


Figure 2. Mean percentage of "good ending" responses as a function of the probe-tone position and the level of musical training of the subjects for the familiar (panel A) and unfamiliar (panel B) material.

response profiles for each condition separately.

Familiar material. The results obtained in this condition are displayed in Panel A of Figure 2. Acceptance rates varied as a function of probe-tone position, $F(11,198) = 61.67$, $MS_e = 3.88$, $p < .001$, in a similar way for musicians and nonmusicians; there was no Musical Expertise effect, $F < 1$, nor was there any interaction between this latter factor and probe-tone position, $F < 1$.

Orthogonal contrasts revealed that all subjects rated the scale tones (with

67.9% acceptance rate, on average) far higher than the nonscale ones (with 21.8% acceptance rate), $F(1,228) = 415.85$, $MS_e = 5.84$, $p < .001$. Acceptance rates among the scale tones were expected to reflect the tonal hierarchy. Comparison of the responses indeed revealed a preference for triad tones (with a mean acceptance rate of 75.1%) over nontriad tones (with a mean acceptance rate of 62.4%), $F(1,228) = 18.6$, $MS_e = 5.84$, $p < .001$. Yet, the preference for the tonic (78.2%) of over the other triad tones (73.6%) did not reach the significance level, $F < 1$.

Because there was a true ending for each fragment (i.e., a specific pitch on which the fragment ended in its original version), memory for the melodies could have contributed to the judgements. Percentage of acceptance of true endings as good completions was indeed extremely high, constituting 96.1% of the responses, and higher than those obtained for any other tone. However, this memory bias cannot account for the observed overall profile because the true endings scanned all the possible scale tones. This was verified statistically: When the responses for the true endings were removed from the data, similar results to those observed for the whole set were obtained.

Acceptance rates for nonscale tones were also found to differ significantly from each other, $F(4,72) = 4.88$, $MS_e = 101.47$, $p < .002$. These fluctuations suggest that among the nonscale tones, some tones did seem to constitute better endings than others. Such differences may reflect the fact that some melodies contained suggestion of different keys than the one considered here as the main key. If correct, the responses should reflect properties of the circle of fifths, as was found by Krumhansl (1990) in similar probe-tone tasks. The circle of fifths can be conceived as a spatial representation of all musical keys (see Cuddy & Thompson, 1992, for a clear description). In that circle, two keys are located one step apart when their tonic are either one fifth — i.e. seven chromatic steps — up (clockwise), or one fifth down (counterclockwise). These keys are said to be “close”, because they share all diatonic tones but one, which is lower or higher by one chromatic step. This tone often serves as an indicator that modulation has just occurred, that is, that the musical key has just smoothly shifted to the key next to it in the circle of fifths. An additional nonscale tone which is very likely to receive a high acceptance rate is the probe-tone distance 3, which serves as an indicator that key has shifted from the major mode to the minor mode, although the *tonic* of the key remains the same. In other words, probe-tone position 3 indicates that the key has moved from, say, C major to C minor. This tone is very important from a music-theoretical point of view, as it is widely used in various “A-B-A” musical forms like Trios or Sonatas. The ratings of the nonscale tones were thus examined on this basis. As can be seen in Figure 2, the data were only partially consistent with such an analysis: The probe-tone position 10 and 3 did indeed receive a high acceptance rate, but probe-tone position did as well ($p < 0.05$ by Tukey *post-hoc* comparisons). Thus, the

nonscale ratings for the familiar tunes cannot be easily accounted for.

Unfamiliar material. The data obtained in the unfamiliar condition are shown in Panel B of Figure 2. The analysis of variance again revealed a large effect of Probe-Tone Position, $F(11,198) = 71.70$, $MS_e = 4.30$, $p < .001$, and no effect of Musical Expertise, $F(1,18) = 1.62$, $MS_e = 22.76$.

Further analyses revealed three significant orthogonal contrasts. Scale tones received a far higher acceptance rate (with a mean acceptance rate of 68.7%) than nonscale tones (with a mean acceptance rate of 19.9%), $F(1,228) = 456.13$, $MS_e = 5.97$, $p < .001$. Triad tones (with 78.9% acceptance rate in average) were preferred over other scale tones (with 61.0% of acceptance rate), $F(1,228) = 36.23$, $MS_e = 5.97$, $p < .001$. Finally, the tonic (with 90% of acceptance rate) was here clearly preferred over the remaining triad tones (with 73.4% of acceptance rate), $F(1,228) = 12.07$, $MS_e = 5.97$, $p < .001$.

As was the case in the familiar condition, acceptance rates for the nonscale tones were also found to be variable, $F(4,76) = 5.27$, $MS_e = 1.76$, $p < .001$. In the present condition, however, a clearer profile emerged: The three most important nonscale tones (i.e., probe-tone positions 6, 3 and 10) were rated higher than probe-tone positions 1 and 8 ($p < .05$ by Tukey *post hoc* comparisons).

To ensure that all these findings were not an artifact of familiarization with the retrograde tune fragments, performance was compared between the two halves of the experiment. No effects of Familiarization, $F(1,18) = 1.91$, $MS_e = 430.55$, nor interaction of this factor with any other factor (all $ps > .05$), were found.

Regression analyses

Given the multidimensionality of the material, it is likely that several aspects of the melodies contributed to the observed response profiles. Among these potential cues are some that are relatively well identified and can be partialled out in regression analyses. Thus, the averaged rating of the 40 subjects provided for each of the 12 probe-tone positions after each of the 7 fragments served as dependent measures in various regression analyses, considering first pitch-based factors and secondly time-based factors. To allow direct comparison between the familiar and the unfamiliar condition, responses that corresponded to the “real ending” of the familiar tunes were dropped from the data.

Pitch-based factors. Three predictor variables related to pitch structure were considered: These include tone repetition, pitch proximity, and tonal weight. Each has previously been found to be a determinant of ending judgements.

Tone repetition has an obvious potential biasing effect on preferences (Butler, 1989); moreover, this factor may also contribute to the building or acquisition of tonal schemas (Krumhansl, 1990). The influence of tone

repetition was evaluated in this study by counting the number of occurrences of the probe tone in the prior context; this varied from 0 to 4.

Because we considered this variable to be a low-level factor, in that it is close to the musical surface, we considered repetition of the *specific pitch*, not of its pitch class¹. Our intention was to evaluate direct short-term memory effects of the probe tone. For example, in familiar tune #3 (see Figure 1), the probe tone D4 was not considered to be a repetition of the initial D3.

Pitch proximity of the last two tones was computed in chromatic steps between the last tone of the context and the probe tone, and varied from 0 to 7. Tone repetition and pitch proximity were contrasted with the contribution of the factor under study, the tonal determinant, designated as the “tonal weight” of the probe tone. Tonal weight was taken from the independent ratings assigned to each degree of the major key by Krumhansl and Kessler (1982). These ratings were made on a 7-point scale, where 7 represented a very good ending to the key-defining context; the values vary from 2.23 to 6.35.

Inter-correlations between predictor variables for the two conditions were calculated. The correlation between repetition and pitch proximity was $r(75) = -.23, p < .05$, for the familiar fragments and $r(82) = -.19, p < .05$, for the unfamiliar fragments. Repetition is found to be highly correlated with tonal weight in both conditions, with $r(75) = .54$ and $r(82) = .60$ for the familiar and unfamiliar conditions, respectively, both $ps < .01$. The correlation between pitch proximity and tonal weight was not significant, $r(75) = .01$ and $r(82) = -.11$ in each condition. This general pattern of correlations supports the notion that tone repetition is closely related to tonal function. Thus, there is a particular need to tease apart their respective contributions to the response profiles obtained in the present study.

Simple correlations (r) between each predictor variable and the mean ratings obtained in each condition are shown in Table 1, mirroring the results of the inter-correlations.

Multiple-regression analysis indicated that, taken together, the three predictor variables accounted for 64% of the variance in the ratings of the familiar material, $F(3,73) = 43.38, p < .001$, and for 65% of the variance in the unfamiliar condition, $F(3,80) = 50.63, p < .001$. The standardized regression coefficients (β weights) shown in Table 1 represent the strength of the relation between the dependent variable (i.e., the ratings) and a predictor when the two other variables are maintained constant. As can be seen, beta

1 Although tones of the same pitch class (that is, tones separated by intervals of one or more octaves) are undoubtedly heard as psychologically close (Krumhansl, 1990, p. 111), experiments have shown that “at some level of processing, octave-equivalent tones are not fully identical” (Krumhansl, 1990, p. 113). Accordingly, octave-equivalent tones cannot be taken as a direct measure of pitch repetition, especially when one considers memory effects.

TABLE 1

Correlations between the pitch-based predictor variables and the mean percentage of "good ending" judgments in the familiar and unfamiliar conditions.

Predictor variable	r	Condition			r	pr ²
		Familiar	pr ²	Unfamiliar		
repetition	.68**	.37**	20.0**	.63**	.24**	9.5**
pitch proximity	-.21	-.134	.42	-.18	-.06	1.23
tonal weight	.71**	.51**	33.1**	.78**	.63**	42.5**

Note. ** represents $p < .01$.

weights (β) for pitch proximity were not significant in any condition. However, coefficients for tone repetition and tonal weight were significant in both conditions.

Partial regression analyses were then performed. The squared value (pr^2) indicates what proportion of the variance in the ratings each predictor variable can explain when the two others have been partialled out. Thus, they represent the unique contribution of each factor to the variance which remains unexplained by the others. As can be seen in Table 1, the best predictor for the ratings in both conditions is undoubtedly the tonal weight of the probe. It added a significant portion to the variance accounted for by the other variables (33% and 42% for the familiar and unfamiliar conditions, respectively). Tone repetition is found to be a lesser predictor than tonal weight but it also brings a significant unique contribution to the variance accounted for in both conditions (20% and 9% for the familiar and unfamiliar conditions, respectively). The unique contribution of pitch proximity was not found to be significant in either condition.

Time-based factors. Although the manipulation of the probe tone was pitch-based here, the context was temporally structured and the task requirements were sufficiently open to allow subjects to integrate temporal features into their decisions. In an attempt to assess the contribution of such factors, two temporally defined variables were considered here: The serial position of the last occurrence of the probe tone, if any, in the prior context, and the metrical stress of the prior occurrences. It will be noted that these temporal features are intrinsic to tone repetition. We conjectured, however, that they may exert distinct influences that may have gone unnoticed in the above regression analysis where tone repetition was computed irrespectively of their temporal position. These time-based factors were tested against the tonal weight of the probe tone, as in the former analysis.

The position of the last occurrence of the probe was measured in terms of tone serial distance. For instance, in a fragment of nine tones, if the probe had been heard for the last time at the fourth position, it was given the score 5 (it was 5 tones away from the last tone). The serial position varied from 1 to 11.

TABLE 2

Correlations between the time-based predictor variables and the mean percentage of "good ending" judgements in the familiar and unfamiliar conditions.

Predictor variable	r	Condition		r	B	pr ²
		Familiar	Unfamiliar			
serial position	.47**	.06	.66	.47**	.18*	6*
metrical stress	.66**	.37**	18.5**	.59**	.153	.8
tonal weight	.71**	.48**	28.7**	.78**	.64**	47**

Note. * represents $p < .05$; ** represents $p < .01$.

Tones at weak metrical locations are less well-remembered than tones at strong metrical locations (Jones, 1976; Jones, Boltz, & Kidd, 1982; Palmer & Krumhansl, 1990). Accordingly, prior occurrences of the probe tone at strong metrical locations may have been more influential on the ending judgement of the probes than tones found at weak locations. Metrical locations are, however, not notated in the transcript, for the tunes are mostly orally transmitted and thus vary in their metrical structure from one transcription to another. We therefore asked three professional musicians to assign time signatures and metrical bars to each of the 14 tunes used in the present study. The metrical interpretation that met the highest agreement among the judges was used here for the attribution of metrical "weight" (see Figure 1 for the transcription of this agreement). The metrical stress for each tone is relative and is coded in terms of the integers 1 to 4 (see Figure 1). The metrical stress for a same-pitch event was summed across all its appearances in the context, amounting to a maximum score of 11.

The outcome of the inter-correlations computed among the three variables suggests that temporal aspects of tone repetition are inter-related, but also may concur with tonal significance. All inter-correlations were found to be significant beyond the .01 level. Correlation between serial distance and metrical stress was $r(75) = .51$ for the familiar condition, and $r(82) = .58$ for the unfamiliar condition. Correlation between serial distance and tonal weight was $r(75) = .45$ for the familiar condition and $r(82) = .33$ for the unfamiliar condition. Finally, tonal weight correlated with metrical stress, with $r(75) = .54$ for the familiar condition and $r(82) = .53$ for the unfamiliar one. It is interesting to note that the correlation between tonal weight and metrical stress is preserved in the unfamiliar tunes. This indicates that in both conditions, tones that are important in the tonal hierarchy are also located at strong metrical positions.

Simple correlations between each predictor variable and the mean ratings are displayed in Table 2. These correlations were also high and were again similar in both conditions. These three predictor variables accounted for 66% of the variance in the familiar condition, $F(3,73) = 46.7$, $p < .001$, and for

67% of the variance in the unfamiliar condition, $F(3,80) = 53.1, p < .001$. Partial regressions indicated that the best predictor for the ratings in both conditions was again the tonal weight of the probe. Metrical stress alone added a significant portion of variance in the familiar condition, but not in the unfamiliar condition. In contrast, serial distance was contributory in the unfamiliar condition, but not in the familiar condition.

DISCUSSION

The primary result of this study is that both musicians and nonmusicians were found to exhibit responses that reflected tonal interpretation of the final tone of tune excerpts, irrespective of their familiarity with the context. Support for this conclusion is based on the obtained tonal hierarchy response profiles; the main evidence, however, is provided by the application of regression procedures designed to disentangle the correlated effects of context-dependent features and tonal knowledge in the judgement of tune completion.

The response profiles displayed by both groups in both conditions were clearly not random but were found to reflect the expected tonal hierarchy. Scale tones were largely preferred over nonscale tones. Among the scale tones, the triad tones received a higher acceptance rate, with the tonic being the most preferred, at least in the unfamiliar condition where the hierarchy among triad tones reached statistical significance.

Similarly, tonal knowledge was found to account for the differential ratings provided for the nonscale tones. For the unfamiliar material, the nonscale tones that were rated as constituting better endings were those that might be expected from an analysis based on the circle of fifths. It is a striking finding that the observed tonal profiles reliably emerged for unfamiliar context, independently of the musical background of the listeners. Given that tonal interpretation of the stimuli was clearer for the unfamiliar material, the question arises as to the role played by familiarity in the present situation.

Familiarity with the musical context definitely played a role here. The ratings for the real endings — that is, the pitch on which the tune ended in its original version — received the highest score, and thereby provide a direct demonstration that the tune representations were indeed accessed and used to perform the task. These true endings even overrode the tonal weight of the final tone, for in most cases they did not correspond to any of the triad tones. Prior knowledge of the tunes therefore had a clear impact on the completion judgements, as far as the true ending is concerned. For the other possible endings, the subjects apparently used the same procedure as with the unknown material. But it was not *quite* the same procedure, for when the true endings were removed from the data, the evidence for the tonal hierarchy still remained less clear than for the unfamiliar material. Thus, it may be tentatively argued that consulting a pre-existing tune representation constrains judgements in such a way that tonal knowledge cannot be expressed fully or freely.

In general, the tonal hierarchy emerged less clearly in the present situation than in the original setting of Krumhansl and Kessler (1982). For instance, as mentioned above, the judgments at the highest level of the hierarchy — i.e., among the triad tones — were not reliably discriminable from each other, in either condition. Several factors can account for these slight divergences from the prototypical tonal hierarchy. An obvious candidate is the dichotomous mode of response that was employed here. Subjects were required to choose between two categories only (“good” or “not”) whereas in the standard situation the response is more fine-grained, usually being couched in terms of 5 or 7 rating scale categories. One could argue that the more subtle the response, the better it can express tonal relations. It is true that the two studies in which a dichotomous response mode was employed (Frankland & Cohen, 1990; Janata & Reisberg, 1988) also reported less clear tonal profiles than those reported by Krumhansl and Kessler. However, several other studies using a 5- or 7-point rating scale have also failed to obtain fine discrimination at this level of the hierarchy (see Krumhansl & Shepard, 1979, for example). Thus, it does not seem likely that the simplified response mode is the sole explanation for the fluctuations observed in the actual profiles.

Much of this variance can be accounted for by the use of real musical fragments. With such stimuli, completion can be determined on a host of different factors. Fortunately, there are conventional means that allow their respective contribution to the completion judgements to be assessed. Regression analyses revealed that the tonal function of the final tone was the primary determinant of the judgements, relative to other potential factors such as tone repetition and pitch proximity. Pitch proximity had no predictive value on ending judgements. Tone repetition, evaluated as an overall index or examined as a function of its serial position or its stress value in the prior occurrence of the probe tone, had only a small effect compared to tonality. Thus, context-dependent knowledge from the specific musical events was found to be less determinant than more abstract, context-free knowledge of tonality. The observed response profiles did not merely mirror the specific context profile (Butler, 1989). Rather, the outcome argues for the determinant role of the abstract tonal knowledge in judging tune completion.

In conclusion, the ecological character of the musical material used in the present experimental setting generally facilitated the application of tonal knowledge to the task. The musical fragments were multidimensional and thus were close to what is usually heard in everyday life. This likely taps more directly, on a more routine basis, the implicit knowledge that all listeners share when listening to tonal music.

This research was supported by a predoctoral fellowship to the first author and by a research grant and a fellowship to the second author, both from the Natural Science and Engineering Research Council of Canada. We are grateful to Luce Beaudet,

Michel Longtin and Luc Bouvrette for assigning metrical interpretations to the tunes. We thank Serge Larochelle and Luc Rousseau for their statistical assistance and advice, and the editor as well as two anonymous reviewers for helpful comments and suggestions on a previous version of this paper.

Correspondence should be addressed to Sylvie Hébert, Département de Psychologie, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal, Québec, H3C 3J7 (e-mail: heberts@ere.umontreal.ca)

References

- Attneave, F., & Olson, R.K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, *84*, 147-166.
- Bharucha, J.J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, *16*, 485-518.
- Boltz, M. (1989). Rhythm and "good endings": Effects of temporal structure on tonality judgments. *Perception & Psychophysics*, *46*, 9-17.
- Boltz, M. (1991). Some structural determinants of melody recall. *Memory & Cognition*, *19*, 239-251.
- 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*, 219-242.
- Cuddy, L.L. (1991). Melodic patterns and tonal structure: Converging evidence. *Psychomusicology*, *10*, 107-126.
- Cuddy, L.L., & Baderstcher, B. (1987). Recovery of tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception & Psychophysics*, *41*, 609-620.
- Cuddy, L.L., & Thompson, W.F. (1992). Asymmetry of perceived key movement in chorale sequences: Converging evidence from a probe-tone analysis. *Psychological Research*, *54*, 51-59.
- Dewar, K.M., Cuddy, L.L., & Mewhort, J.K. (1977). Recognition memory for single tones with and without context. *Journal of Experimental Psychology*, *1*, 60-67.
- Dowling, W.J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341-354.
- Francès, R. (1958). *La perception de la musique*. Paris: J. Vrin.
- Frankland, B.W., & Cohen, A.J. (1990). Expectancy profiles generated by major scales: Group differences in ratings and reaction time. *Psychomusicology*, *9*, 175-183.
- Janata, P., & Reisberg, D. (1988). Response-time measures as a means of exploring tonal hierarchies. *Music Perception*, *6*, 161-171.
- Jones, M.R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*, 323-335.
- 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 & Psychophysics*, 32, 211-218.
- Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Krumhansl, C.L., & Keil, F.C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory & Cognition*, 10, 243-251.
- Krumhansl, C.L., & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 89, 334-368.
- Krumhansl, C.L., & Shepard, R.N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 579-594.
- Lerdahl, F. (1988). Tonal pitch space. *Music perception*, 5, 315-350.
- Meyer, L.B. (1956). *Emotions and meaning in music*. Chicago: University Chicago Press.
- Palmer, C., & Krumhansl, C.L. (1990). Mental representation for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 728-741.
- Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, 10, 21-56.
- Peretz, I., & Morais, J. (1989). Music and modularity. *Contemporary Music Review*, 4, 277-291.
- Peretz, I., Babai, M., Lussier, I., Hébert, S., & Gagnon, L. (1995). Corpus d'extraits musicaux: indices de familiarité, d'âge d'acquisition et d'évocation verbale. *Revue canadienne de psychologie expérimentale*, 49, 211-239.
- Serafine, M.L. (1983). Cognition in music. *Cognition*, 15, 119-183.
- Speer, J.R., & Meeks, P.U. (1985). School children's perception of pitch in music. *Psychomusicology*, 5, 49-56.

Date of acceptance: July 31, 1993

Sommaire

Les règles tonales dans l'écoute musicale

Les études portant sur l'utilisation des règles tonales dans l'écoute musicale rapportent souvent que les auditeurs non musiciens, contrairement aux musiciens, ont de la difficulté à mettre ces principes en vigueur lorsqu'ils effectuent une tâche en laboratoire. Dans une tâche de jugement de l'achèvement d'une séquence, les réponses des non-musiciens sont plutôt influencées par des facteurs contextuels relatifs aux stimuli utilisés, tels que la proximité en hauteur par rapport à la dernière note de la séquence musicale, que par la fonction tonale de la note finale de la séquence. La présente étude vise à déterminer si les non-musiciens manifestent mieux leur savoir tonal lorsque de vraies mélodies plutôt que les stimuli traditionnels tels que les gammes ou les suites d'accords sont utilisés. Deux groupes d'auditeurs — des musiciens et des non-musiciens — ont jugé l'acceptabilité de chacune des 12 notes de la gamme chromatique comme note finale d'airs familiers et non familiers. Les airs non familiers étaient l'inversion en miroir des airs familiers, en termes de hauteurs et de durées des notes. Les résultats indiquent que les réponses des musiciens et des non-musiciens sont guidées par les relations tonales pour les deux types de matériel. En effet, les notes de la gamme diatonique ont été jugées comme étant plus acceptables que les notes ne faisant pas partie de la gamme diatonique (premier niveau de hiérarchie tonale). Aussi, parmi les notes de la gamme diatonique, les notes formant l'accord de tonique ont reçu plus d'acceptation que les autres notes (deuxième niveau de hiérarchie tonale). Finalement, parmi les notes de l'accord de tonique, la tonique a reçu un plus haut degré d'acceptation, mais pour le matériel non familier seulement (troisième niveau de hiérarchie tonale). Afin de départager l'influence de la connaissance tonale de celle de facteurs spécifiques aux stimuli utilisés, des analyses de régression ont été effectuées. Ces dernières révèlent clairement que les facteurs contextuels tels que la proximité en hauteur des deux dernières notes, le nombre de répétitions de note à l'intérieur des airs, ou encore la position temporelle du dernier son ont peu d'influence sur les patrons de réponses observés. La contribution respective de tous ces facteurs est en effet minime par rapport à l'importance de la note finale dans la hiérarchie tonale, confirmant que les auditeurs ont fondé leur jugement sur la base de leurs connaissances abstraites de la structure musicale. Ainsi, le caractère écologique des stimuli employés ici semble avoir favorisé l'utilisation des connaissances implicites que tous les auditeurs partagent lors de l'écoute musicale.