

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Exploring the Influence of Pitch Proximity on Listener's Melodic Expectations

J. Fernando Anta

Faculty of Fine Arts, National University of La Plata

Author Note

J. Fernando Anta, Department of Music, Faculty of Fine Arts, National University of La Plata.

This research was supported by grants from the National University of La Plata and from the National Council for Scientific and Technical Research (CONICET) of Argentina. The author is grateful to Dr. Mayumi Adachi for critical reading of an earlier draft of the manuscript, and to Dr. Annabel J. Cohen, Dr. Bradley W. Frankland, Dr. E. Glenn Schellenberg, and one anonymous reviewer for their helpful comments during the review process.

Correspondence concerning this article should be addressed to J. Fernando Anta, Department of Music, Faculty of Fine Arts, National University of La Plata, Diagonal 78 N° 680, La Plata (1900), Argentina. E-mail: fernandoanta@fba.unlp.edu.ar

1 *Abstract*

2 Two studies, one correlational and one meta-analytic, were conducted to explore whether
3 pitch proximity influences listeners' melodic expectations about pitch direction and tonality.
4 Study 1 used a probe-tone task. Specifically, listeners heard fragments of tonal melodies
5 ending on intervals of 1 or 2 semitones, and rated how well individual tones (the probe-tones)
6 continued each fragment. Regression analyses showed that listeners expected probe-tones to
7 be proximate in pitch to the last tone they heard in the fragments, but not to the penultimate
8 one, since there was no evidence of expectations for a change in pitch direction. However,
9 listeners expected probe-tones to be proximate to (at least one of) the other tones they had
10 heard in the fragments, since there was evidence of expectations for a melodic movement
11 toward the bulk of the fragments' pitch distribution. In addition, the most stable probe-tones
12 in the key of the fragments were more expected than the least stable ones only when they
13 were proximate to the tones presented in the fragments. The results of Study 1 were replicated
14 and extended in Study 2, in which a meta-analysis of data reported in Schellenberg (1996,
15 Experiment 1) was performed. These data had been collected using the same probe-tone task
16 as in Study 1, but different melodic fragments; the fragments ended on intervals of 2 or 3
17 semitones. Together, the present findings suggest, first, that when small intervals occur in a
18 melody, pitch proximity has only a global influence on expectations about pitch direction; and
19 second, that pitch proximity constrains the influence of tonality on melodic expectation.

20 *Keywords:* Melodic expectation – pitch proximity – pitch direction – tonality –
21 statistical and heuristic learning – conceptualization.

22

23

24

1 **Pitch Proximity and Melodic Expectation**

2 The generation of expectations about what events will occur next in a musical piece
3 and when these will appear has long been thought to be a critical aspect of music listening.
4 Basically, the idea is that expectations that prove to be correct facilitate music understanding,
5 or vice versa, and that the interplay between expectations' negation and resolution regulates
6 the emotional content of musical experience (Huron, 2006; Meyer, 1956; Narmour, 1990).
7 During the last decades, this idea has stimulated a great deal of theoretical and empirical
8 research, which has been particularly fruitful in the domain of pitch-related melodic
9 expectation. Research in this domain revealed that to a large degree listener's expectations
10 can be accounted for by three factors, namely pitch proximity, pitch direction, and tonality
11 (see, for example, Bharucha, 1996; Carlsen, 1981; Huron, 2006; Larson, 2004; Lerdahl, 1996,
12 2001; Margulis, 2005; Meyer, 1956, 1973; Narmour, 1990; Schellenberg, 1997; Temperley,
13 2008). The present research further investigated these factors by exploring how and to what
14 extent pitch proximity influences expectations about pitch direction and tonality.

15 The influence of pitch proximity refers to the effect of distance in pitch between
16 sequentially presented tones on the listener's expectations. It is understood as a gestalt-like
17 auditory principle, perhaps the most basic one, regulating what tones listeners expect to occur
18 next in a melody in order to be grouped with the tones heard previously (Larson, 1997;
19 Margulis, 2005; Meyer, 1956; Narmour, 1990; see also Bregman, 1990). In line with this,
20 there is strong evidence that listeners expect the next tone in a melody to be proximate in
21 pitch to the last tone they heard; all things being equal, the nearer in pitch the ensuing tone is
22 to the last one, the more highly it is expected (Carlsen, 1981; Cuddy & Lunney, 1995; Eerola,
23 2004; Krumhansl, 1995; Larson, 2004; Schellenberg, 1996). For example, given the melody
24 shown in Figure 1, at the beginning of measure 4 (see the '?' mark in the figure) the
25 expectation for Bb4 would be stronger than the expectation for B4, because the former is

1 nearer to A4 (the last tone in the melody) than the latter; the most expected tone would be A4,
2 the unison. Following Krumhansl (1995) and Schellenberg (1996), this pattern of expectation
3 was coded into a numerical predictor variable called “Pitch Proximity” (see Figure 1), in
4 which one unit is subtracted whenever distance between the current tone of the melody (e.g.,
5 A4 at the end of measure 3) and the next tone (e.g., the ‘possible next tones’ at the beginning
6 of measure 4) increases by one semitone.¹

7 It is important to notice that pitch proximity, as described so far, is a local or first-
8 order factor regulating melodic expectation, in the sense that it refers to the “expectedness” of
9 tones solely based on (distance to) the last tone heard. However, there is evidence that
10 listeners also expect the next tone in a melody to be proximate in pitch to the penultimate tone
11 heard (Krumhansl, 1995; Schellenberg, 1996, 1997). This means that pitch proximity
12 regulates melodic expectation as a slightly more global or second-order factor too
13 (Schellenberg et al., 2002).² Moreover, it also means that listeners’ expectations of pitch
14 direction emerge, at least in part, from the second-order influence of pitch proximity:
15 Inasmuch as listeners expect tones to be proximate to the penultimate tone heard, they expect
16 a return to the earlier pitch register and, therefore, a reversal in direction (Narmour, 1990;
17 Schellenberg et al., 2002). Given the melody shown in Figure 1, for example, listeners would
18 expect a reversal of direction toward B4 more strongly than a continuation toward G4,
19 because the former is nearer to the penultimate tone (Bb4) than the latter. Following
20 Schellenberg (1997), this pattern of expectation was coded into a numerical predictor variable
21 called “Pitch Reversal” (see Figure 1), in which ‘1.5’ is assigned to tones that reverse pitch
22 direction (upward to downward or vice versa) and land proximate (± 2 semitones) to the
23 penultimate tone in the melody, and ‘0’ is assigned to the remaining tones.

24 However, two issues need to be addressed before accepting the second-order influence
25 of pitch proximity on the expectation of pitch direction. The first is that this type of

1 expectation may arise from other factors. For example, listeners may envisage pitch direction
2 based on interval size, expecting large intervals to be followed by a reversal and small
3 intervals to be followed by a continuation of direction (Aarden, 2003; Krumhansl, 1995;
4 Thompson et al., 1997; von Hippel, 2002). Although the second-order influence of pitch
5 proximity may explain the former of these expectations (Schellenberg et al., 2002; see also
6 Narmour, 1990), it cannot explain the latter one. Alternatively, listeners may simply expect
7 melodic movements toward the most stable tones of the key, that is to say, toward the tones of
8 the tonic chord (Larson, 2004; Povel, 1996; see also Bharucha, 1996). Thus, given the melody
9 of Figure 1, listeners would expect a reversal toward Bb4 not because of the second-order
10 influence of pitch proximity, but because Bb4 is the tonic of the underlying key, Bb Major.
11 Following Krumhansl (1995) and Schellenberg (1996), the expectation for continuation of
12 direction that would emerge from small intervals was coded into a categorical predictor
13 variable called “Registral Direction”, set as ‘1’ for tones satisfying such expectation and ‘0’
14 otherwise, whereas expectations for tonal stability were coded into a numerical predictor
15 variable called “Tonal Hierarchy”, quantified using the index proposed by Krumhansl and
16 Kessler (1982; see Krumhansl, 1990, p. 30). Figure 1 illustrates how Registral Direction and
17 Tonal Hierarchy were quantified.³

18 The second issue that needs to be addressed is that, in some cases, the second-order
19 influence of pitch proximity may also be thought of as its first-order influence. Specifically,
20 when small intervals occur in a melody, to hypothesize that listeners expect the next tone to
21 be proximate “to the penultimate tone they heard” or “to the last tone they heard” is, at least
22 partially, the same thing, because tones proximate to the penultimate tone are proximate to the
23 last tone too. In the melody shown in Figure 1, for example, B4 (as a possible next tone) will
24 be proximate to both Bb4 and A4, the last two tones in the melody. Indeed, the predictions
25 encoded in Pitch Reversal and Pitch Proximity shown in Figure 1, respectively representing

1 the second-order and first-order influence of pitch proximity, are significantly correlated ($r =$
2 $.43, n = 25, p = .03$), which highlights their conceptual overlap. This conceptual overlap
3 finally suggests that, when small intervals occur in a melody, the second-order influence of
4 pitch proximity on melodic expectation could be weakened or even nullified by its first-order
5 influence, simply because the former and the latter are almost the same. This hypothesis was
6 examined in the present research, as will be detailed later.

7 In any case, if one accepts that pitch proximity has not only a first-order but also a
8 second-order influence on expectation, thus promoting expectations of pitch reversal, then one
9 should ask whether its influence extends further backward, involving even higher orders of
10 melodic processing, and whether in this way it affects expectations of direction too. In simple
11 terms, listeners might expect tones to be proximate not only to the last or the second to last
12 tone they heard, but also to the third to last one, to the fourth to last one, and so on. Then, they
13 might simply expect the melody to move toward the bulk of its range (i.e., toward the center
14 of its pitch distribution), to reach a tone proximate to (at least one of) the tones that occurred
15 previously. If so, given the melody of Figure 1, listeners would strongly expect a reversal of
16 direction, because A4 is one extreme of the range, and tones that continue in the same
17 direction (as the last interval, Bb4-A4) will move away from all of the previous tones. This
18 idea, in turn, fits well with probabilistic models of melodic expectation, which hypothesize
19 that listeners expect melodies to remain within their range (Temperley, 2008; von Hippel &
20 Huron, 2000). However, from the present perspective listeners would expect a reversal in
21 direction because of a global influence of pitch proximity and not, as probabilistic models
22 claim, because of listeners' sensitivity to the "regression to the mean (or central pitch)"
23 tendency that melodies show. In fact, it may be argued that this tendency is a numerical
24 artifact (Huron, 2006). To test if pitch proximity has a backward global influence on listeners'
25 expectations of pitch direction, a categorical predictor variable called "Range Direction" (see

1 Figure 1) was created, set as ‘1’ for tones occurring toward the bulk of the range of the
2 melody, and ‘0’ otherwise; the range’s “bulk” was defined as its central pitch (e.g., C#5 in the
3 melody shown in Figure 1).

4 A final issue addressed in the present research was whether pitch proximity affects
5 listeners’ tonal expectations. In this respect, previous studies have shown that listeners’
6 expectations for the most stable tones of the key are stronger than those for the remaining
7 tones (Krumhansl, 1995; Larson, 2004; Marmel et al., 2010; Schellenberg, 1996, 1997;
8 Schellenberg et al., 2002; see also Boltz, 1989, 1993). In showing this, it was assumed that
9 tones separated by one or more octaves convey the same level of (tonal) stability and,
10 consequently, that they are equally expected to occur next in a melody regardless of the
11 distance in pitch between them and the tones previously heard. That is to say, it was assumed
12 that tonal expectations emerge independently of pitch proximity. In the Tonal Hierarchy
13 predictor variable illustrated in Figure 1, this is reflected in the fact that possible next tones an
14 octave away (e.g., Bb3 and Bb4) receive the same value. However, there are at least three
15 reasons to doubt this assumption. The first is simply that, as far as the author knows, whether
16 or not this is the case has not yet been directly tested. The second is that octave equivalence
17 effects may not operate directly in the processing of melodic sequences (Deutsch, 1969, 1983;
18 Deutsch & Boulanger, 1984), particularly when listeners without extensive musical training
19 are examined (Dowling & Hollombe, 1977; Kallman, 1982; Sergeant, 1983). This raises the
20 possibility that melodic patterns having the same pitch-classes but in different octaves (e.g.
21 A4-Bb4 and A4-Bb5, in Bb major) are perceived as having different levels of stability.

22 Finally, the third reason is that the ‘most stable’ tones of the key are perceived that
23 way in part because they are near in pitch to a previously occurring unstable tone; otherwise,
24 they do not provide the ‘resolution’ or the anchoring effect that characterizes them (Bharucha,
25 1984a, 1996; Bigand et al., 1996). This suggests that tonality’s influence on melodic

1 expectation is constrained by pitch proximity: Listeners would expect not only the occurrence
2 of stable tones but also the anchoring of unstable tones and, therefore, that the former will be
3 near in pitch to the latter. Interestingly, melodic anchoring can be delayed by an intervening
4 tone between the unstable tone and its anchor (Bharucha, 1984a), which suggests that pitch
5 proximity also imposes a second-order constraint on tonal expectations. If this is the case,
6 then one should also ask whether the constraint imposed by pitch proximity extends further
7 backward, involving even higher orders of melodic processing. If so, listeners would expect
8 stable tones to be proximate not only to the last or the second to last tone, but also to the other
9 tones they heard in a given context. That is to say, listeners would simply expect the
10 occurrence of the nearest stable tones, those that would serve as adequate anchors. To test this
11 hypothesis, two numerical predictor variables were created, “Tonal Hierarchy-8” and “Tonal
12 Hierarchy-6/5”. They were quantified by retaining the values of Tonal Hierarchy within the
13 range of an octave and a sixth-or-a-fifth, respectively, around the tones of the melody; given
14 that other (more distant) anchors would not be expected, all other values were set at zero.
15 Figure 1 illustrates how Tonal Hierarchy-8 and Tonal Hierarchy-6/5 were quantified.⁴

16 To summarize, previous studies show that pitch proximity has both a first-order and a
17 second-order influence on melodic expectation, and that due to its second-order influence it
18 affects listener’s expectations of pitch direction—the listener expecting a reversal in direction.
19 However, it is not clear whether the first-order influence and the second-order influence of
20 pitch proximity can be differentiated when small intervals occur in a melody. In addition, it
21 has not been determined whether the influence of pitch proximity on expectations of direction
22 extends further backward, involving even higher orders of melodic processing, nor whether
23 the influence of tonality is independent of pitch proximity. To address these issues, the
24 following two studies were conducted.

1 **Study 1: Probe-tone ratings of melodic continuity**

2 The purpose of the present study was to investigate the influence of pitch proximity on
3 listener's melodic expectations about pitch direction and tonality. In order to achieve this
4 purpose, a probe-tone task was used. In a probe-tone task, a musical context is presented to
5 listeners and, next, a single tone (referred to as the 'target' or 'probe') is played; then,
6 listeners are asked to rate the tone according to some criterion. The overall idea is that a
7 highly expected (probe-)tone will be given a high rating, and vice versa. Probe-tone tasks
8 have been widely used to examine listeners' melodic expectations (e.g., Cuddy & Lunney,
9 1995; Krumhansl, 1995; Schellenberg, 1996, 1997; Schellenberg et al., 2002; see also Pearce
10 & Wiggins, 2006); in fact, the one used in the present study was largely based on those used
11 in Schellenberg's (1996) work—which will be revisited in Study 2. It has to be acknowledged,
12 however, that probe-tone tasks have been subject to various criticisms (see, for example,
13 Aarden, 2003; Huron, 2006; see also Butler, 1990), some of which will be considered later, in
14 the Results and Discussion section. This notwithstanding, for the purpose of the present study
15 a probe-tone task was particularly useful, because it allows one to systematically control the
16 distance in pitch between the tones of the contexts and the tones used as probes. Additionally,
17 it has the advantage of providing a detailed picture of the issue being studied when listeners
18 judge several probes for a set of fixed contexts (Huron, 2006), as was the case here.

19 **Method**

20 **Participants.** Twenty-eight undergraduate students of the National University of La
21 Plata participated in the study. One half of the participants (5 females and 9 males; 22.4 –
22 27.3 years old, $M = 24.3$) were students of music (hereafter referred to as “musicians”) who
23 had completed at least four years of musical training (i.e., music theory, ear training, and
24 instrumental performance) within the past six years. None of them was a vocal student.
25 Therefore, it was assumed that they would not be familiar with the musical excerpts used as

1 stimuli (see Materials)—which could have affected the way in which they performed the
2 experimental task (see Procedure); informal post-test interviews supported this assumption.
3 The other half of the participants (again, 5 females and 9 males; 22.7 – 28.4 years old, $M =$
4 25.8) were students of natural sciences (hereafter referred to as “non-musicians”) who had
5 never studied music formally. All participants reported having normal hearing and were
6 unaware of the purposes of the study.

7 **Materials.** Eight fragments of Western tonal melodies were used as contexts for
8 testing the hypotheses outlined above (see Appendix A); one of them was the melodic
9 fragment shown in Figure 1. The fragments were taken from the vocal line of *Lieder* (i.e., art
10 songs for voice and piano accompaniment) by Franz P. Schubert (1797–1828): Fragment 1
11 was from Schubert’s Opus 25 (*Die schöne Müllerin*, D. 795), whereas fragments 2 to 8 were
12 from Schubert’s Opus 89 (*Winterreise*, D. 911). It is worth mentioning that Schubert’s
13 musical pieces, from which opuses 25 and 89 stand out, may be seen as highly representative
14 of the Western tonal repertoire regarding both their tonal (or ‘harmonic’) and voice-leading
15 (or ‘melodic’) structure (Piston 1941/1959; Salzer 1952/1982). The fragments had between 8
16 and 12 tones and started at the beginning but ended before the completion of a melodic group.
17 Four of the fragments were in a major key and the other four were in a minor key. Specifically,
18 based on key signatures and the presence/absence of alterations, for fragments 1 to 8, the
19 assigned keys were: Bb major, Ab major, A minor, C major, A minor, G minor, Eb major, and Bb
20 minor. Key assignment was further assessed with three well-known methods: first, the key-finding
21 algorithm of Longuet-Higgins and Steedman (1971), which supported the keys assigned to
22 fragments 1, 2, 3, 5, 6, and 7; second, the ‘rare-interval’ theory of Brown and Butler (1981; see also
23 Browne, 1981), which supported the keys assigned to fragments 1, 3, 5, and 7; and third, the key-
24 finding algorithm of Krumhansl and Schmuckler (see Krumhansl, 1990), which supported the keys

1 assigned to fragments 2, 4, 7, and 8. As may be seen, these analyses suggested that overall
2 fragments would be heard in the assigned key.

3 The fragments were chosen such that, firstly, the final interval (i.e., between the last
4 two tones) was small. Specifically, half of the fragments ended on an interval of 1 semitone,
5 and the other half ended on an interval of 2 semitones (the smallest intervals used in Western
6 tonal music). Secondly, compared with the penultimate tone of the fragment, the last tone
7 occurred in a weaker metrical position and was less stable in a tonal sense (as measured by the
8 tonal hierarchy index from Krumhansl and Kessler (1982)). Thus, the final interval of each
9 fragment was “unclosed” (i.e., promoted a sense of non-closure or continuation), according to
10 theories of music perception (e.g., Lerdahl & Jackendoff, 1983; see also Lerdahl, 2001) and
11 expectation (e.g., Narmour, 1990). Thirdly, the last tone was one extreme of the melodic range
12 (as in the fragment shown in Figure 1), or was nearer to one extreme than the other. Lastly,
13 the fragments were chosen such that the final interval was moving away from the center of the
14 range, or going toward it (four fragments each); for example, in the fragment shown in Figure
15 1 the interval Bb4-A4 moves away from the center of the range (C#5), and the interval A4-
16 Bb4 would have moved toward it.

17 For each melodic fragment, 15 probe-tones were used. Probe-tones consisted of the
18 diatonic tones in a two-octave range centered on the final tone of the melodic fragment given
19 as context; for example, the probe-tones used for the fragment shown in Figure 1 were A3,
20 Bb3, C4, D4, Eb4, F4, G4, A4, Bb4, C5, D5, Eb5, F5, G5 and A5. Each probe-tone had the
21 same duration and temporal location (following the fragments’ last tone) as the tone that
22 actually followed in the *Lieder* from where fragments were taken. Fragments and probe-tones
23 were presented with a synthesized piano timbre at the *tempi* suggested in the original scores.
24 The dynamic level of stimuli was adjusted by participants according to their preference.

1 **Apparatus.** Initially, stimuli were created as musical instrument digital interface
2 (MIDI) files using notation software (Finale) installed on a laptop computer (Hewlett-Packard
3 G60-123CL). Next, stimuli were converted into sound tracks with *Garritan Instruments for*
4 *Finale*[®] set to the “*Steinway Piano*” timbre, and recorded as MP3 files on the computer. The
5 computer also controlled the presentation of the stimuli. Finally, stimuli were reproduced
6 stereophonically (JVC/MIX-J10 loudspeakers) in a sound-attenuated room.

7 **Procedure.** The guidelines of the procedure were borrowed from the literature (e.g.,
8 Cuddy & Lunney, 1995; Krumhansl, 1995; Schellenberg, 1996; Schellenberg et al., 2002).
9 Specifically, participants were told that they would hear fragments of melodies that started at
10 the beginning but ended before the completion of a musical phrase. Then, they were told that,
11 in successive presentations of a given fragment, a single tone would be added to the end of the
12 fragment. Finally, participants were asked to rate how well each added tone (i.e., each probe-
13 tone) continued the fragment on a scale from 1 (*extremely bad*) to 10 (*extremely good*). In line
14 with the literature, participants were also told that each melodic fragment would still sound
15 incomplete even with the tone added, and that they should not rate how well the added tone
16 “completed” the fragment but rather how well it “continued” the fragment. Finally, they were
17 urged to use the entire rating scale, reserving the endpoints for extreme cases.

18 The participants were tested individually. Each participant performed a practice
19 session and, next, a test session. Each test session consisted of eight blocks of trials, one for
20 each of the melodic fragments described in the section on Materials. Each block consisted of a
21 melodic fragment presented twice and, from its third presentation, the fragment plus a single
22 probe-tone added until all probe-tones were heard; a silence of about 5-6 seconds preceded
23 each fragment presentation, during which listeners rated (since the third presentation) each
24 probe-tone. The practice session was equivalent in all respects to the test session, except that
25 it was shorter; only two fragments were used, also extracted from *Lieder* by Schubert (see

1 Appendix A). After the practice session, listeners were left alone and the test session began.
2 The participants used the computer keyboard to initiate each block of trials and to record their
3 responses. Within each session, the order of blocks as well as the order of probe-tones in each
4 block was randomized separately for each participant. Each practice session lasted
5 approximately 15 minutes, and each test session lasted approximately 50 minutes; during the
6 test session, participants were allowed to take short breaks as needed.

7 **Results and Discussion**

8 For the analyses presented herein, only data collected in the test sessions were
9 considered: 120 data points (15 probe-tones for each of the 8 melodic fragments) for each
10 participant. Initially, the agreement between participants' judgments was examined by means
11 of correlation analyses. The average intersubject correlations within groups were $r = .416$
12 (range: $r = .183 - r = .708$) for musicians, and $r = .513$ (range: $r = .180 - r = .792$) for non-
13 musicians (all correlations $n = 120, p < .05$). Because of the significant within-group
14 correlations, the average ratings of each group of listeners were used in subsequent analyses.
15 These analyses, in turn, showed that musicians and non-musicians' judgments could not be
16 explained by the same set of predictor variables (see below). Therefore, their ratings were not
17 averaged.

18 Next, the correlations between the average ratings of musicians and non-musicians
19 and the predictor variables presented above (for a summary, see Figure 1) were computed; the
20 correlations between predictor variables were also computed, using the values corresponding
21 to the 120 probe-tones the participants rated. Table 1 shows the results of these analyses. Two
22 of them are particularly important. The first is that the correlation between musicians' average
23 ratings and Tonal Hierarchy (TH) was not significant. Given that only diatonic (and not
24 chromatic) tones were used as probes, this result appeared to suggest that the experimental
25 design had not been sensitive enough to capture the influence of tonality on musicians'

1 judgements. However, the significant correlation between non-musicians' average ratings and
2 TH argued against this interpretation. More critically, then, it suggested that musicians'
3 judgments had not been influenced by tonality, which was in contrast to what had been
4 reported in previous studies (e.g., Krumhansl, 1995; Larson, 2004; Schellenberg, 1996, 1997;
5 Thompson et al., 1997; see also Krumhansl et al., 2000).

6 To explore this point in more detail, the orientation toward a dominant setting of
7 musicians' tonal expectations was assessed, specifically examining whether musicians
8 expected the next tone of the fragments to be part of the dominant chord (instead of being part
9 of the tonic chord, as assumed in TH). The tonal orientation toward V/I was chosen because
10 both music theory (Piston, 1941/1959; Salzer, 1952/1982) and psychology (Krumhansl &
11 Kessler, 1982; Lerdahl, 2001) suggest that the dominant chord possesses the second-greatest
12 level of stability (next to the tonic chord) in a given key. A new predictor variable was then
13 designed, referred to as Tonal Hierarchy [Dominant] (TH[D]), in which the tonal hierarchy
14 index was adjusted to a dominant setting. For the fragment shown in Figure 1, for example, in
15 TH[D] the dominant tone (F) received the highest value, followed by the other tones of the
16 dominant chord (A and C), and finally by the other diatonic tones (G, Bb, D, and Eb). As
17 shown in Table 1 (right-side columns), TH[D] was significantly correlated with the average
18 ratings of musicians (but not with the average ratings of non-musicians), suggesting that their
19 judgments had also been influenced by tonality. Finally, two other predictor variables were
20 designed to examine whether pitch proximity had constrained the influence of tonality on
21 musicians' judgments, Tonal Hierarchy [Dominant]-8 (TH[D]-8) and Tonal Hierarchy
22 [Dominant]-6/5 (TH[D]-6/5). As their tonic-oriented counterparts, TH[D]-8 and TH[D]-6/5
23 were quantified by retaining the values of TH[D] within the range of an octave and a sixth-or-
24 a-fifth, respectively, around the last tones heard. For example, for the fragment shown in
25 Figure 1 values were retained from F5 to F4 in TH[D]-8, and from F5 to A4 in TH[D]-6/5.

1 The second important result shown in Table 1 is the redundancy or multicollinearity
2 that existed between the predictor variables tested, which is reflected in the fact that all of
3 them were positively correlated with at least one of the remaining predictors. This result was
4 not surprising, since most of the predictor variables were designed here to code the influence
5 of pitch proximity on melodic expectation, or already coded this influence (as is the case of
6 Pitch Reversal). Further, TH-8 and TH-6/5 (or TH[D]-8 and TH[D]-6/5) also coded listeners'
7 tonal expectations, as did TH (or TH[D]). However, multicollinearity deserved special
8 attention because it suggested that in a multivariate analysis of listener's judgments some of
9 the collinear predictors could be discarded without sacrificing explanatory power. Moreover,
10 multicollinearity could reduce any individual variable's predictive power by the extent to
11 which it was associated with the other predictor variables (Ho, 2006).

12 Taking multicollinearity into account, hierarchical regression analyses were used to
13 examine the multivariate structure of participants' data. Hierarchical regression allows one to
14 deal with multicollinearity and, in addition, is particularly safe when the influence of several
15 factors on a given outcome is initially explored (Ho, 2006), as was the case in this study.
16 Briefly, in a hierarchical regression model, predictor variables are entered into the equation in
17 blocks or steps in such a way that the most important predictors (based on logical or
18 theoretical considerations) are entered first. Then the less important predictors are entered and
19 evaluated in terms of what they add to the explanation above and beyond that afforded by the
20 predictors already entered. In addition, at each step of the analyses the *stepwise* variable
21 selection procedure was used. This means that the predictor variables were entered one at a
22 time if they significantly improved the ability of the model to explain the data, and deleted at
23 any step if they no longer contributed significantly. Thus, for each data set a rather
24 parsimonious model was obtained.

1 In sum, in the first step of the analyses, two explanatory variables were entered to
2 control for extraneous sources of variance. These variables were Frequency of occurrence
3 (FO) and Recency (Re), which coded the number of times and the order in which each tone
4 occurred in each fragment. Previous studies show that these variables may also influence the
5 formation of melodic expectations (Krumhansl, 2000; Oram & Cuddy, 1995; Tillman et al.,
6 2000). In fact, it has been suggested that when a probe-tone task is used, listeners' ratings
7 could simply reflect the effects of short-term memory for tones presented in the contexts
8 (Butler 1990). Therefore, by controlling the influence of FO and Re it was evaluated whether
9 the influence of pitch proximity and tonality on melodic expectation could be distinguished
10 from the effects of short-term memory. Next, in the second step of the analyses, Pitch
11 Proximity (PP) was entered, the predictor variable that received the strongest support from
12 previous studies. In the third step, Pitch Reversal (PR), Registral Direction (RD) and Range
13 Direction (RaD) were entered, that is, the various predictor variables related to listeners'
14 expectations about pitch direction. Finally, in the fourth step the predictor variables entered
15 were those related to listeners' tonal expectations, TH[D], TH[D]-8 and TH[D]-6/5 when
16 analyzing musicians' data and TH, TH-8 and TH-6/5 when analyzing non-musicians' data.
17 The results of the analyses are summarized in Table 2.

18 As may be seen in Table 2 (see Step 4), the models that provided the best fit to the
19 data for each training group are almost equal. Although Re was significant only for musicians
20 and FO only for non-musicians, the predictors entered in the final model for each group were
21 PP, RaD, and the predictor that coded tonal expectations in a range of an octave (TH[D]-8 and
22 TH-8 for musicians and non-musicians, respectively). That is to say, both groups of listeners
23 showed a first-order influence and a global influence of pitch proximity on their expectations
24 (as coded in PP and RaD, respectively), and also a limited influence of tonality (as coded in
25 TH[D]-8 or TH-8). In contrast, PR, RD, and the remaining tonal predictors (TH[D], TH[D]-

1 6/5, TH and TH-6/5) were not reliable, and were excluded from the final models. This means,
2 first, that the second-order influence of pitch proximity on listeners' expectations (as coded in
3 PR) was negligible; second, that the influence of the direction of the last interval heard (as
4 coded in RD) was also negligible; and third, that the influence of tonality was not spread
5 throughout the register (as coded in TH[D] or TH) nor was it limited to a too narrow range (as
6 coded in TH[D]-6/5 or TH-6/5).

7 Two other results shown in Table 2 are noteworthy. First, as reflected in the
8 corresponding β (standardized Beta) value, the association between Re and musicians' ratings
9 was negative (see Step 4). However, based on previous studies (e.g. Tillman et al., 2000;
10 Toiviainen & Krumhansl, 2003; see also Oram & Cuddy, 1995), it was expected to be
11 positive (with listeners rating higher the probe-tones more recently heard). The negative
12 association then suggests that musicians did not judge probe-tones according to standard tonal
13 criteria which, in turn, lends further support to the idea that they did not expect the occurrence
14 of tones of the tonic chord (but that of tones of the dominant chord).⁵ Second, the sr^2 (squared
15 semipartial correlation) value of (i.e., the proportion of variance explained uniquely by) each
16 predictor was low, but the values of tolerance were relatively safe. Briefly, these values
17 indicate the percentage of variance in each predictor that cannot be accounted for by the other
18 predictors and, then, whether multicollinearity between predictors is problematic; recall that,
19 to some degree, the predictors entered in Step 4 were collinear (see Table 1). Roughly
20 speaking, tolerance values below .20 indicate a potential problem (Menard, 2002; see also Ho,
21 2006): as shown in the table, all the tolerance values were above this limit.

22 Subsequent analyses were conducted to further explore whether pitch proximity had
23 exerted a second-order influence on listeners' judgments, and whether the influence of
24 tonality had actually been constrained in range. Regarding the first issue, it could be that pitch
25 proximity had exerted a second-order influence, but that the predictor variable used to test this

1 possibility, Pitch Reversal, had not grasped it. Therefore, two alternative predictor variables
2 were created: Pitch Proximity[minus]1 (PP-1) and Change of Direction (CoD). PP-1 was
3 designed as PP was, but coding expectations for tones proximate to the penultimate tone
4 heard; for the fragment shown in Figure 1, for example, the zero value of PP-1 was centered
5 on Bb4 (and not on A4, as in PP). CoD simply coded expectations for reversals in direction;
6 given the melody shown in Figure 1, for example, probe-tones higher than A4 were coded as
7 '1', and the remaining probe-tones as '0'. Regarding the second issue, two interaction terms
8 were created to test whether instead of 'being constrained', tonality's influence had decreased
9 gradually across the register: PPxTH(D) and PPxTH. Then, the regression analyses were
10 conducted again but this time PP-1 and CoD were also entered in Step 3, and only TH(D) or
11 TH plus the corresponding interaction term (PPxTH(D) and PPxTH when analysing
12 musicians' and non-musicians' data, respectively) were entered in Step 4. These analyses
13 showed that PP-1 and CoD made no significant contribution to the models tested, either for
14 musicians' judgments ($\beta = .15$ and $.06$, respectively; $ps > .05$) or non-musicians' judgments (β
15 $= .13$ and $.05$, respectively; $ps > .05$). The interaction terms were not significant either ($\beta = -$
16 $.11$ and $.21$, respectively; $ps > .05$).

17 Finally, a series of analyses was conducted in order to further assess the validity of the
18 regression models previously estimated using the stepwise procedure (see Table 2, Step 4).
19 Specifically, these analyses were intended to examine, first, whether the models performed
20 equally well across melodic fragments, and second, whether they performed actually better
21 than (or at least similar to) the other models one could construct with the predictor variables
22 that had been excluded from the analyses. With these purposes in mind, the judgments of
23 continuation provided by each participant for each fragment (i.e., each set of 15 ratings the
24 participants gave within each block of trials) were successively regressed on five different
25 two-step models in which the explanatory variables FO and Re were entered first, in Step 1,

1 and PP plus the ‘target’ predictor variables (i.e., those coding expectations about pitch
2 direction and tonality) were entered next, in Step 2. It is worth mentioning that in these
3 analyses the *forced entry* (instead of the stepwise) variable selection procedure was used,
4 which means that all the predictor variables in a given model were allowed to enter the
5 regression equation; therefore, differences between the fits of the models would depend only
6 on the predictors used. However, the sr^2 values of the predictor variables entered in Step 2
7 were subtracted from the overall fit (R^2) of each model if they were found to be negatively
8 correlated with the data (recall that all predictors were expected to have positive coefficients):
9 had this ‘correction’ not been made, two or more models would have shown a similar (or
10 indeed, the same) fit even if they were based on the opposite predictions (e.g., if they included
11 PD or RD, in the case of Fragment 1). The models tested and the results of these analyses are
12 depicted in Figure 2.

13 As may be seen in Figure 2, the performance of Models 4, which contained the (target)
14 predictors previously selected with the stepwise procedure, was roughly constant across
15 fragments, particularly in the case of non-musicians’ data. Clearly, musicians’ data showed
16 more dispersion. However, in six of the eight fragments (fragments 1, 3, 5, 6, 7, and 8) the fit
17 of Model 4 varied within the range of .52 – .66 mean R^2 (i.e., a range of only 14% of
18 explained data), with a grand mean of $R^2 = .58$ (SE = .02). This suggests that, overall, the
19 models obtained with the stepwise selection procedure are relatively resistant to extraneous
20 contextual factors (i.e., factors other than those related to pitch proximity and tonality).
21 Perhaps more importantly, Figure 2 also shows that the mean R^2 of Models 4 was similar to or
22 even higher than the mean R^2 of the other models tested, particularly in the case of non-
23 musicians’ data. (It is worth mentioning that for fragments 3, 4, 7, and 8, the mean R^2 of
24 Models 2 is similar to that of Models 4—and indeed matches that of Models 3—simply because
25 RD and RaD encoded the same predictions.)

1 In order to examine this point more closely, the R^2 values obtained from each model
2 for each participant across the various melodic fragments were converted to z -scores and,
3 finally, entered into a 2 x 5 mixed-design analysis of variance (ANOVA) with Training
4 (musicians or non-musicians) as between-subjects factor and Model (models 1 to 5) as within-
5 subjects factor. Mauchly's test indicated that the assumption of sphericity had been violated
6 for the effect of Model ($\chi^2(9) = 380.45, p < .001$); therefore degrees of freedom were
7 corrected using the Greenhouse–Geisser estimates of sphericity ($\epsilon = .54$). This correction
8 revealed a significant main effect of Model ($F(2.17, 482.64) = 36.91, p = .001$), with Models
9 3 to 5 performing similarly to each other but outperforming Models 1 and 2, and a significant
10 interaction between Model and Training ($F(2.17, 482.64) = 6.39, p = .001$), with the
11 difference between Models 3 to 5 and Models 1 and 2 being larger in non-musicians' data.
12 This means, first, that overall the expectations about pitch direction coded in RaD (entered in
13 Models 3 to 5) were more relevant than those coded in PR and RD (entered in Models 1 and
14 2, respectively); second, that the tonal expectations coded in TH[D] or TH (entered in models
15 number 3) but not in TH[D]-8 or TH-8 (entered in Models number 4) could be dropped
16 without loss of explanatory power (since models 3 and 4 performed equally well); and third,
17 that the relevance of the expectations of direction coded in RaD and the 'uselessness' of the
18 tonal expectations dropped from the analyses were even more marked in non-musicians' data.

19 To summarize, in agreement with previous studies (e.g., Carlsen, 1981; Cuddy &
20 Lunney, 1995; Krumhansl, 1995; Larson, 2004; Schellenberg, 1997), Study 1 showed a first-
21 order influence of pitch proximity on listeners' melodic expectations (as coded in PP).
22 However, in contrast to previous studies, it was found that when small intervals occur in a
23 melody, the second-order influence of pitch proximity (as coded in PR, PP-1, and CoD) is
24 negligible; the influence of interval direction (as coded in RD) was also found to be
25 negligible. In addition, it was found that pitch proximity has a backward global influence on

1 melodic expectation, thus affecting expectations of direction (as coded in RaD), and that the
2 influence of tonality is limited to a narrow pitch range (as coded in TH[D]-8 and TH-8 for
3 musicians and non-musicians, respectively). Overall, then, the results of Study 1 indicated
4 that the way in which pitch proximity influences melodic expectation is more complex than
5 previously suggested. Nevertheless, this needed to be confirmed, given the exploratory nature
6 of the study. Therefore, a second study was undertaken.

7

8 **Study 2: a Meta-Analysis of Schellenberg (1996, Experiment 1)**

9 The purpose of Study 2 was to examine whether the results of Study 1 could be
10 replicated in a different sample, derived from different musical materials. As mentioned
11 above, given the exploratory nature of Study 1, the results needed to be confirmed. However,
12 this was also necessary due to methodological reasons. Specifically, when a hierarchical
13 regression analysis is performed, as in Study 1, the final model may not be suitable for
14 explaining or predicting a new set of (comparable) data. This would be particularly the case
15 when the stepwise variable selection procedure is used, as was done in Study 1. In a stepwise
16 regression, decisions about which variables are entered or removed from the model are solely
17 based on statistical criteria, as explained above. Therefore, a stepwise regression model tends
18 to capitalize on chance variation in the sample at hand (Ho, 2006). To overcome these
19 problems, it can be assessed whether the model generalizes beyond the initial sample: if a
20 model can be generalized, then it must be capable of explaining the same response variable
21 from the same set of predictor variables in a different sample. With this idea in mind, in Study
22 2, a sample of continuation judgments reported by Schellenberg (1996) was reanalyzed to see
23 whether the model that fits them better was one of the models validated in Study 1.

24 **Method**

1 The sample used in Study 2 was taken from a database reported in Schellenberg (1996,
2 Experiment 1: Appendix A, fragments 1 to 4). In this work, Schellenberg conducted three
3 probe-tone experiments to investigate listeners' melodic expectations. In the first experiment,
4 the musical contexts were 8 fragments of tonal melodies taken from British folk songs. (In the
5 second and third experiments, the contexts were fragments of atonal and pentatonic
6 melodies.) All fragments ended on an unclosed interval; four of them ended on a small
7 interval of 2 or 3 semitones, while the other four ended on a large interval of 9 or 10
8 semitones. It is worth mentioning that, following Narmour (1990), Schellenberg (1996)
9 considered melodic intervals of 3 semitones as small; therefore, the same was done here.
10 Probe-tones consisted of the 15 diatonic tones in a two-octave range centered on the last tone
11 of the melodic context. The participants, 20 listeners with moderate or limited musical
12 training (10 listeners each), were asked to rate how well additional single tones (i.e., the
13 probe-tones) continued the melodic fragments on a scale from 1 (*extremely bad*) to 7
14 (*extremely good*). The experiment yielded 120 continuation ratings for each participant (8
15 fragments x 15 probe-tones).

16 Next, Schellenberg (1996) correlated the data from each listener with those from every
17 other listener in order to evaluate consistency across participants. All pairwise correlations
18 were statistically significant, which warranted the averaging of the ratings across participants
19 and training levels. Finally, Schellenberg examined whether participants' average ratings
20 reflected the influence of pitch proximity, pitch direction, and tonality. Briefly, a first-order
21 and a second-order influence of pitch proximity on participants' judgments was observed, as
22 well as an influence of tonality (see also Schellenberg, 1997). However, the data collected
23 from the fragments ending on small intervals were not separately analyzed; they were
24 analyzed together with the data collected from the fragments ending on large intervals.
25 Therefore, it was not directly tested whether pitch proximity has a second-order influence on

1 melodic expectation when small intervals occur in a melody, nor whether small intervals
2 generate expectations of melodic direction. In addition, it was not investigated whether the
3 influence of pitch proximity extends further backward, involving even higher orders of
4 melodic processing, nor whether the influence of tonality is constrained by pitch proximity. In
5 order to evaluate these alternatives, the present study focused on the analysis of the average
6 continuation ratings collected from the four fragments ending on small intervals in
7 Experiment 1 of Schellenberg (1996) (see Appendix B). That is to say, 60 average data points
8 were analyzed (4 fragments x 15 probe-tones), as described below.

9 **Results and Discussion**

10 First, the correlations between the average continuation ratings from Schellenberg
11 (1996) and the predictor variables from Study 1 were computed; the correlations between
12 predictor variables were also computed, using the values corresponding to the 60 probe-tones
13 the participants rated in Schellenberg's experiment (when the fragments ending on small
14 intervals were given). Table 3 shows the results of these analyses. As can be seen in the table,
15 the average ratings from Schellenberg's experiment were not significantly correlated with
16 Tonal Hierarchy (TH), nor with Tonal Hierarchy[Dominant] (TH[D]). However, in line with
17 Schellenberg's report, the correlation coefficients of the tonic-oriented tonal predictors were
18 larger than those of the dominant-oriented tonal predictors. Therefore, only TH, TH-8 and
19 TH-6/5 were retained for subsequent analyses. In addition, Table 3 also shows that, as in
20 Study 1, there was a great deal of multicollinearity between many of the predictor variables.
21 Then, hierarchical regression analyses were conducted to examine the multivariate structure
22 of Schellenberg's data.

23 The regression analyses were carried out in the same way as in Study 1. Initially,
24 average continuation ratings were analyzed using a four-step hierarchical regression in which
25 all independent variables were successively entered into the regression equation. In the first

1 step, the Frequency of occurrence (FO) and the Recency (Re) of the tones in the musical
2 contexts were entered as explanatory variables. Next, in the second step, Pitch Proximity (PP)
3 was entered into the equation. In the third step, Pitch Reversal (PR), Registral Direction (RD)
4 and Range Direction (RaD) were entered, that is, the predictors related to listeners'
5 expectations about pitch direction. Finally, in the fourth step the predictors entered were those
6 related to listeners' tonal expectations, in this case, TH, TH-8, and TH-6/5. In addition, in
7 each step of the regression the stepwise variable selection procedure was used, as was done in
8 Study1. The results of these analyses are summarized in Table 4.

9 Table 4 shows that the statistically significant predictor variables in the final model
10 (see Step 4) were PP, RaD, and TH-6/5; when TH-6/5 was entered, Re was no longer
11 significant. As the reader may note, this result closely resembles those observed in Study 1
12 (see Table 2, Step 4). Again, support was found for a model of melodic expectation in which,
13 when small intervals occur in a melody, pitch proximity has both a local and a global (rather
14 than a second-order) influence, and tonality has an influence in a limited pitch range. Note
15 also that, again, the sr^2 value of each predictor was low, but the tolerance values were well
16 above the limit of .20. Further, as in Study 1, the existence of a second-order influence of
17 pitch proximity and the possibility of a graded decline in tonality's influence were explored
18 by regressing the data on a model in which Pitch Proximity[minus]1 (PP-1) and Change of
19 Direction (CoD) were also entered in Step 3, and only TH and its interaction term with PP
20 were entered in Step 4 and, again, PP-1 and CoD made no significant contribution to the
21 model ($\beta = -.14$ and $.11$, respectively; $ps > .05$), nor did the interaction term ($\beta = -.32$, $p =$
22 $.31$).

23 Notwithstanding the convergence between the results from this second study and those
24 of Study 1, there is an important difference. Specifically, in Schellenberg's data the influence
25 of tonality was limited to an even narrower range, of a sixth or a fifth; in Study 1, its influence

1 was observed in a range of an octave. This raises the possibility that the constraint imposed by
2 pitch proximity on tonal expectations is stronger than suggested by Study 1. In the present
3 study, however, it seems possible and even likely that tonality's influence has also been
4 constrained by another factor, specifically, by tones' recency. In this line, note that, overall,
5 the melodic range of the fragments from Schellenberg (1996) and that of the fragments used
6 in Study 1 are similar. However, close to the end (e.g., during the last three or four beats,
7 based on time signature), the pitch register of the former (a perfect fifth or less) tend to be
8 narrower than that of the latter (which rather suddenly traverse a major sixth—Fragment 8—or a
9 minor seventh—fragments 2 and 3—). Thus, listeners from Schellenberg's experiment could
10 have had more limited tonal expectations not because the constraint of pitch proximity is
11 stronger than previously thought, but because the tones they heard most recently were closer
12 to each other in pitch. Future research should explore this possibility.

13 Finally, a set of analyses was conducted in order to further assess the validity of the
14 regression model previously estimated using the stepwise procedure (see Table 4, Step 4), as
15 was done in Study 1. Specifically, the average data from each melodic fragment were
16 regressed on five different two-step models in which the explanatory variables FO and Re
17 were entered first, in Step 1, and PP plus the target predictor variables (i.e., those coding
18 expectations about pitch direction and tonality) were entered next, in Step 2. As in Study 1, in
19 these analyses the forced entry method was used, and the sr^2 of the predictor variables entered
20 in Step 2 were subtracted from the overall fit of each model if they were found to be
21 negatively associated with the data. The models tested and the results of the analyses are
22 depicted in Figure 3. As may be seen, Model 5, which contained the (target) predictors
23 previously selected by the stepwise procedure, performed similarly or even better than the
24 other models tested. In addition, it performed almost equally well across fragments, which

1 supports the idea that a model of melodic expectation based on pitch proximity and tonality is
2 relatively resistant to other contextual factors.

3 Figure 3 also shows that Model 5 outperformed models 1 to 4 in the case of fragments
4 1 and 2, but not in the case of fragments 3 and 4, where Model 3 achieved the best fit. (In
5 fragments 3 and 4, the R^2 of Models 2 and 3 is the same because RD and RaD encoded the
6 same predictions). Given that Models 5 and 3 differed only in the inclusion of TH-6/5 or TH,
7 respectively, these results suggested that, at least for fragments 3 and 4, the influence of
8 tonality on listeners' expectations had not been limited in range, as indicated by the results of
9 the stepwise regression analysis (in which TH-6/5 instead of TH was selected; see Table 4,
10 Step 4). A final set of regression analyses was conducted to evaluate this possibility.
11 Specifically, the average continuation ratings from fragments 3 and 4 were regressed on all
12 the predictor variables of Model 5, the residuals of the model were calculated (i.e., the
13 differences between observed and predicted scores) and, finally, regressed on TH. In this way,
14 it was examined whether TH could account for a significant proportion of the variance in the
15 data not accounted for by TH-6/5, nor by either of the remaining variables entered in Model 5.
16 The result of these analyses showed that TH did not account for a significant proportion of the
17 variance in the residuals of Model 5 neither when data from Fragment 3 ($R^2 = .17$, $F(1,13) =$
18 2.59 , $p = .13$) nor when data from Fragment 4 ($R^2 = .15$, $F(1,13) = 2.36$, $p = .15$) were
19 considered. This confirmed that the hypotheses (expectations) coded in TH but not in TH6/5
20 could be removed from the analysis without a significant loss of predictive accuracy.

21 In sum, the results of Study 2 largely replicated the results of Study 1. In both studies
22 it was found, first, that pitch proximity has a first-order influence and a backward global
23 influence on listeners' melodic expectations (as coded in PP and RaD, respectively); second,
24 that the second-order influence of pitch proximity (as coded in PR, PP-1, and CoD) and the
25 influence of intervals' direction (as coded in RD) are negligible when small intervals occur in

1 a melody; and third, that the influence of tonality is limited to a narrow pitch range (as coded
2 in TH-6/5 in the case of Study 2, and in TH-8 or TH[D]-8 in the case of Study 1). In addition,
3 the results of Study 2 revealed that the second-order influence of pitch proximity on melodic
4 expectation becomes perceptually relevant only when relatively large melodic intervals occur
5 in a melody. In this regard, recall that two of the melodic fragments from Schellenberg's
6 experiment (fragments 3 and 4) ended on an interval of 3 semitones, and the second-order
7 influence of pitch proximity was not yet detectable.

8

9

General Discussion

10 The present research explored how and to what extent pitch proximity affects
11 listeners' melodic expectations about pitch direction and tonality. In this respect, previous
12 studies (e.g., Cuddy & Lunney, 1995; Krumhansl, 1995; Schellenberg, 1996, 1997;
13 Schellenberg et al., 2002; Thompson et al., 1997) reported that pitch proximity has both a
14 first-order and a second-order influence on melodic expectation, and that due to its second-
15 order influence listeners expect a reversal of pitch direction. It was also reported that the
16 influence of tonality spreads across octaves, independently of the distance in pitch between
17 tones (see also Tillmann et al., 2000; Toiviainen & Krumhansl, 2003). The two studies
18 reported here lend further support to the first-order influence of pitch proximity, but suggest
19 that its second-order influence is negligible when small intervals occur in a melody.
20 Notwithstanding this, evidence was found for a backward, more global influence of pitch
21 proximity, which would prompt listeners to expect melodic movements toward the center of
22 the melody's range. In addition, it was found that the most stable tones of the key are more
23 expected than the least stable ones only when upcoming tones are near in pitch, which
24 suggests that tonality's influence on melodic expectation is constrained by pitch proximity.

1 Do the present studies constitute fair tests of the second-order influence of pitch
2 proximity on melodic expectation? Is the influence of tonality always constrained by pitch
3 proximity as was the case here? Concerning the first of these questions, the answer is clearly
4 yes. It has to be borne in mind that the second-order influence of pitch proximity would
5 simply depend on the distance in pitch between the last two tones heard in a melody
6 (Narmour, 1990; Schellenberg, 1997; see also footnote 2). There is some evidence that it
7 becomes stronger as listeners increase in age or acquire more exposure to music, but it was
8 found to be fully operative in adult listeners (Schellenberg et al., 2002), such as those tested
9 here. Therefore, it should be observable even when small intervals occur in a melody, as was
10 the case at the end of the melodic fragments used as context.

11 Concerning the second question, it cannot be answered so readily. For instance,
12 tonality's influence might be wider (or less constrained) if the range of a melody is larger than
13 those of the fragments used in the present studies. It might be wider also when a large interval
14 occurs in a melody, or when the melody quickly traverse a large register (as discussed in
15 Study 2), since distant anchors would be (almost) simultaneously activated. As may be noted,
16 these possibilities capitalize on the context-dependent relationship that appears to exist
17 between pitch proximity and tonality during melodic expectation. However, tonality's
18 influence might be wider if not only expectations for diatonic tones but also for chromatic
19 tones are examined; recall that only the former were examined here. Given that chromatic
20 tones are particularly unstable in a tonal context (Krumhansl & Shepard, 1979; Krumhansl &
21 Kessler, 1982), one would expect a preference for diatonic over chromatic tones even when
22 upcoming tones in a melody are distant in pitch. It has to be pointed out, however, that the
23 way listeners represent the tonal hierarchy of tones in Western tonal music is more complex
24 than a "diatonic versus chromatic" two-level structure (Krumhansl, 1990; Lerdahl, 2001). In
25 fact, strictly speaking, the tonal hierarchy index provided by Krumhansl and Kessler (1982)

1 has twelve levels (or stability values) for each mode (for the major mode, see Figure 1), out of
2 which seven (i.e., those concerning diatonic tones) were evaluated here. Notwithstanding this,
3 and in order to be cautious, it seems reasonable to conclude that tonality's influence on
4 melodic expectation is constrained by pitch proximity when diatonic tones are given. Future
5 research should explore whether this situation holds when chromatic tones are given as well.

6 How does the present work relate to previous works on melodic expectation? How
7 does it contribute to improve our understanding about the subject? Most of all, it seems to
8 highlight the importance of learning in the formation of melodic expectations, as current
9 research increasingly does (e.g., Huron, 2006; Pearce & Wiggins, 2006; Tillmann et al., 2000;
10 see also Meyer, 1956, 1973). On the one hand, learning may account for the different tonal
11 orientation of musicians and non-musicians' judgments observed in Study 1. Concerning this,
12 it must be noted that when musicians talk about tonal music, concepts such as 'continuation'
13 (or non-closure) and 'completion' (or closure) refer theoretically to the tonal functions of
14 dominant and tonic, respectively (see, for example, Piston 1941/1959; Salzer 1952/1982; see
15 also Stein & Spillman, 1996). Thus, it seems that when asked to judge 'continuation' (as
16 opposed to 'completion'), musicians relied on their musical theoretical background and
17 expected the occurrence of tones belonging to the dominant chord (instead of tones belonging
18 to the tonic chord); non-musicians would have expected the occurrence of tones belonging to
19 the tonic chord (which may be seen as the 'by default' tonal expectations) because they had
20 no such theoretical background. This implies that tonality's influence on expectation is
21 partially regulated by the conceptual framework listeners use to understand music, which
22 arises in large part due to music training. Certainly, this hypothesis needs to be explored
23 further. However, there is some evidence supporting it (e.g., Krumhansl et al., 1987; see also
24 Huron, 2006, pp. 168-172), and it is also consistent with previous studies that argue for the

1 importance of concepts, abstraction, and generalization in music perception and expectation
2 (e.g., Meyer, 1956; see also DeBellis, 1995; Zbikowski, 2002).⁶

3 On the other hand, learning may provide a comprehensive explanation of the results
4 observed in both of the studies reported in this article. Concerning this, research in music
5 theory and cognition shows that there are many organizational regularities in music, and
6 particularly in melody, that listeners might learn and that might prompt them to form the
7 expectations observed here. For example, in a wide variety of melodies, adjacent tones tend to
8 be proximate in pitch, which may explain why the participants expected the next tone of the
9 fragments to be proximate in pitch to the last tone they heard (Huron, 2001; Thompson &
10 Stainton, 1998; von Hippel, 2000). Melodies also tend to remain within a narrow pitch range,
11 which may explain why the participants expected melodic movements toward the bulk of the
12 fragments' pitch distribution (Temperley, 2008; von Hippel, 2000; see also von Hippel &
13 Huron, 2000). Indeed, there is evidence that also in the melodies of Schubert's *Lieder* and
14 British folk songs (from which fragments used in studies 1 and 2 were borrowed) adjacent
15 and non-adjacent tones tend to be near in pitch (Huron, 2001; Temperley, 2008; von Hippel,
16 2000) which, in the context of the present argument, is highly suggestive. Finally, in tonal
17 melodies the less stable tones of the key are usually 'resolved' or anchored by steps (e.g.
18 Piston 1941 [1959]; see also Bharucha, 1996), which may explain why the participants
19 expected only the occurrence of the nearest anchors. Thus, the learning of melodic regularities
20 may account for the first-order influence of pitch proximity on melodic expectation, its
21 backward global influence, and the limited influence of tonality.

22 It is, however, important to clarify the idea that this research highlights the importance
23 of learning in melodic expectation. As mentioned above, current research suggests that
24 listeners' expectations may be accounted for in terms of the learning of organizational
25 regularities in music. This leads one to conclude that musical learning is statistical in nature,

1 and that statistical learning is the basis for music expectation: Listeners would learn the
2 probabilistic structure of music, and would expect the most probable events in a given context
3 (Eerola, 2004; Pearce & Wiggins, 2006; Temperley, 2008; Tillmann et al., 2000). However, it
4 has also been suggested that listeners tend to overgeneralize their musical knowledge, and that
5 when generating expectations they often rely on heuristics or ‘rules of thumb’ that are helpful
6 in predicting most musical situations, but that also are mere approximations (basically,
7 simplifications) of the probabilistic structure of music (Huron, 2006; see also von Hippel,
8 2002). This implies that musical learning is heuristic in nature, and that heuristic learning is
9 the basis for music expectation: Listeners would not expect what is more probable to occur
10 next in a given musical context, but what is simpler and more serviceable given their previous
11 experience of similar situations. Although the present findings do not constitute a decisive
12 proof of this issue, they are more consistent with heuristic than with statistical learning. At
13 least three reasons may be given to support this statement.

14 First, the first-order influence of pitch proximity, its second-order influence (when
15 large intervals occur), and its backward global influence may be seen as heuristics to
16 anticipate upcoming tones. Although adjacent and non-adjacent tones in a melody tend to be
17 near in pitch, how near they are varies from melody to melody. Thus, it seems reasonable to
18 conclude that listeners do not calculate how proximate upcoming tones should be each time
19 they hear a melody, but simply expect adjacent and non-adjacent tones to be proximate (see
20 Huron 2006; von Hippel, 2002). Second, the influence of tonality, as captured by the tonal
21 hierarchy index used here, may be seen as a heuristic too. Although the sense of tonality
22 largely emerges from the statistical distribution of pitch-classes in tonal pieces (Krumhansl,
23 1990; Temperley & Marvin, 2008; Tillmann et al., 2000), not all (nor even two) pieces in a
24 given key share the same probabilistic pitch-class structure. Thus, it seems reasonable to
25 conclude that listeners do not calculate the distribution of pitch-classes each time they hear a

1 tonal melody, but simply expect tones according to their schematic knowledge of what is
2 common in tonal music. The tonal hierarchy index from Krumhansl and Kessler (1982) would
3 capture such schematic knowledge (Bharucha, 1984b; Krumhansl, 1990). Moreover, the
4 limited influence of tonality may also be seen as a heuristic: Even if listeners are able to
5 understand any tonal relationship that might exist between more or less distant tones, when an
6 unstable tone occurs in a melody they would still expect only the occurrence of the nearest
7 anchors, because that would be a simpler and more serviceable prediction about how tonal
8 melodies do behave.

9 Finally, the third reason is that pitch proximity and tonality (i.e., the index from
10 Krumhansl and Kessler (1982)) were reliable predictors of listeners' judgments even after
11 controlling for the effects of the frequency of occurrence and the recency of tones, which are
12 the hallmarks of statistical learning (see, Krumhansl, 1990; Oram & Cuddy, 1995; Temperley,
13 2007; Tillmann et al., 2000). This suggests that, to some degree, melodic expectations are
14 based on heuristics regardless of whether or not they are based on statistics. There is some
15 evidence, however, that this would not be the case if further statistics are considered.
16 Specifically, Pearce and Wiggins (2004, 2006) reported that statistical models of melodic
17 expectation in which 19 predictors (plus some interactions) were entered almost entirely
18 subsumed the functions of pitch proximity and tonality (as coded in Pitch Proximity, Pitch
19 Reversal, and Tonal Hierarchy) in accounting for previously collected behavioral data; part of
20 the data re-analyzed by Pearce and Wiggins was taken from Schellenberg (1996, Experiment
21 1), as was done here in Study 2. However, it must be noted, first, that the function of pitch
22 proximity and tonality was only *almost* entirely subsumed; second, that the role of
23 overgeneralization in melodic expectation could be found to be larger if further heuristics are
24 considered too; and third, that in order to evaluate whether heuristics guide melodic
25 expectation, they must be properly defined. Indeed, Pearce and Wiggins assumed that pitch

1 proximity has a second-order influence even when small intervals occur in a melody, and that
2 tonality's influence is relevant even when upcoming tones are distant in pitch, whereas the
3 present findings suggest that neither of these assumptions is necessarily correct.

4 In conclusion, the present findings suggest that the influence of pitch proximity on
5 melodic expectation extends further backward than might previously have been thought.

6 Although pitch proximity would not exert a second-order influence when small intervals
7 occur in a melody, it would still influence expectation at higher, more global levels of melodic
8 processing. This backward, more global influence of pitch proximity would prompt listeners
9 to expect melodic movements toward the bulk of the melody's range, and would also prompt
10 them to expect only the nearest anchors. From a more general perspective, finally, the present
11 findings suggest that the cognitive machinery underlying the formation of melodic
12 expectations is highly intricate: overall a given factor might be critical (e.g., pitch proximity),
13 but in some situations it might be irrelevant (e.g., when small intervals are heard); the role of
14 one factor (e.g., tonality) might be constrained by some other factor (e.g., by pitch proximity);
15 expectations might be generated in a statistical but also in a heuristic manner; and, at the end
16 of the process, conceptual knowledge might alter the expectations generated from musical
17 data (e.g., from the tonal content of music). Certainly, if this is an adequate description of how
18 melodic expectation works, the structure of such complex machinery needs to be explored
19 further. It is hoped that this research will aid future efforts in this regard.

20

21

22

23

24

References

- 1
- 2 Aarden, B. (2003). Dynamic melodic expectancy. Unpublished doctoral dissertation, Ohio
3 State University, Columbus, USA.
- 4 Aldwell, E., Schachter, C., & Cadwallader, A. (2011). *Harmony & Voice Leading* (4th ed.).
5 Boston, MA: Schirmer/Cengage Learning.
- 6 Bharucha, J. J. (1984a). Anchoring effects in music: the resolution of dissonance. *Cognitive*
7 *Psychology*, *16* (4), 485-518.
- 8 Bharucha, J. J. (1984b). Event hierarchies, tonal hierarchies, and assimilation: A reply to
9 Deutsch and Dowling. *Journal of Experimental Psychology: General*, *113*(3), 421-425.
- 10 Bharucha, J. J. (1996). Melodic anchoring. *Music Perception*, *13* (3), 383-400.
- 11 Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord
12 sequences: The influence of harmonic function, sensory dissonance, horizontal motion,
13 and musical training. *Perception & Psychophysics*, *58*, 125-141.
- 14 Boltz, M. (1989). Rhythm and "good endings": Effects of temporal structure on tonality
15 judgments. *Perception & Psychophysics*, *46* (1), 9-7.
- 16 Boltz, M. (1993). The generation of temporal and melodic expectancies during musical
17 listening. *Perception & Psychophysics*, *53* (6), 585-600.
- 18 Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*.
19 Cambridge, MA: MIT Press.
- 20 Brown, H., & Butler, D. (1981). Diatonic trichords as minimal tonal cue-cells. *In Theory*
21 *Only*, *5* (6 & 7), 39-55.
- 22 Browne, R. (1981). Tonal implications of the diatonic set. *In Theory Only*, *5* (6 & 7), 3-21.
- 23 Butler, D. (1990). Tonal hierarchies and rare intervals in music cognition-response. *Music*
24 *Perception*, *7*, 325-338.

- 1 Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology*,
2 *1*, 12-29.
- 3 Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals:
4 Perceptual judgments of continuity. *Perception and Psychophysics*, *57* (4), 451-462.
- 5 DeBellis, M. (1995). *Music and Conceptualization*. Cambridge: Cambridge University Press.
- 6 Deutsch, D. (1969). Music recognition. *Psychological Review*, *76* (3), 300-307.
- 7 Deutsch, D. (1983). Octave equivalence in the processing of tonal sequences. *Journal of the*
8 *Acoustical Society of America*, *74*, 10.
- 9 Deutsch, D., & Boulanger, R. C. (1984). Octave equivalence and the immediate recall of pitch
10 sequences. *Music Perception*, *2* (1), 40-51.
- 11 Dowling, W. J., & Hollombe, A. W. (1977). The perception of melodies distorted by splitting
12 into several octaves: Effects of increasing proximity and melodic contour. *Perception*
13 *and Psychophysics*, *21*, 60-64.
- 14 Eerola, T. (2004). Data-driven influences on melodic expectancy: Continuations in north
15 sami yoiks rated by south African traditional healers. In S. D. Lipscomb, R. Ashley, R.
16 O. Gjerdingen and P. Webster (Eds.), *Proceedings of the 8th International Conference*
17 *on Music Perception and Cognition* (pp. 83-87), Evanston, IL, USA.
- 18 Ho, R. (2006). *Handbook of univariate and multivariate data analysis and interpretation with*
19 *SPSS*. Boca Raton, FL: Chapman & Hall/CRC.
- 20 Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual
21 principles. *Music Perception*, *19*, 1-64.
- 22 Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge,
23 MA: MIT Press.
- 24 Kallman, H. J. (1982). Octave equivalence as measured by similarity ratings. *Perception &*
25 *Psychophysics*, *32* (1), 37-49.

- 1 Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford
2 University Press.
- 3 Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects.
4 *Music Theory Spectrum, 17*, 53–90.
- 5 Krumhansl, C. L. (2000). Tonality induction: A statistical approach applied cross-culturally.
6 *Music Perception, 17*(4), 461-479.
- 7 Krumhansl, C. L., & Kessler, E. (1982). Tracing the dynamic changes in perceived tonal
8 organization in a spatial representation of musical keys. *Psychological Review, 89*, 334-
9 368.
- 10 Krumhansl, C. L., Sandell, G. J., & Sergeant, D. C. (1987). The perception of tone hierarchies
11 and mirror forms in twelve-tone serial music. *Music Perception, 5* (1), 31-78.
- 12 Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions
13 within a diatonic context. *Journal of Experimental Psychology, 5* (4), 579-594.
- 14 Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J.
15 (2000). Cross-cultural music cognition: cognitive methodology applied to North Sami
16 yoiks. *Cognition, 76* (1), 13-58.
- 17 Larson, S. (1997). The problem of prolongation in tonal music: Terminology, perception, and
18 expressive meaning. *Journal of Music Theory, 41* (1), 101-136.
- 19 Larson, S. (2004). Musical forces and melodic expectations: Comparing computer models and
20 experimental results. *Music Perception, 21* (4), 457-498.
- 21 Lerdahl, F. (1996). Calculating tonal tension. *Music Perception, 13* (3), 319-363.
- 22 Lerdahl, F. (2001). *Tonal pitch space*. Oxford: Oxford University Press.
- 23 Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge MA:
24 MIT Press.

- 1 Longuet-Higgins, H. C., & Steedman, M. J. (1971). On interpreting bach. *Machine*
2 *Intelligence*, 6, 221–241.
- 3 Margulis, E. H. (2005). A model of melodic expectation. *Music Perception*, 22 (4), 663-714.
- 4 Marmel, F., Tillmann, B., & Delbé, C. (2010). Priming in melody perception: Tracking down
5 the strength of cognitive expectations. *Journal of Experimental Psychology: Human*
6 *Perception and Performance*, 36 (4), 1016–1028.
- 7 Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.). Thousand Oaks, CA: Sage
8 Publications.
- 9 Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- 10 Meyer, L. B. (1973). *Explaining music: Essays and explorations*. Chicago: University of Chicago
11 Press.
- 12 Narmour, E. (1990). *The analysis and cognition of basic melodic structures*. Chicago:
13 University of Chicago Press.
- 14 Oram, N., & Cuddy, L. L. (1995). Responsiveness of Western adults to pitch-distributional
15 information in melodic sequences. *Psychological Research*, 57 (2), 103–118.
- 16 Pearce, M. T., & Wiggins, G. A. (2004). Rethinking gestalt influences on melodic
17 expectancy. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen and P. Webster (Eds.),
18 *Proceedings of the 8th International Conference on Music Perception and Cognition*
19 (pp. 367-371), Evanston, IL, USA.
- 20 Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: The influence of context and
21 learning. *Music Perception*, 23 (5), 377-405.
- 22 Piston, W. (1941/1959). *Harmony*. London: Victor Gollancz.
- 23 Povel, D.-J. (1996). Exploring the elementary harmonic forces in the tonal system.
24 *Psychological Research*, 58 (4), 274–283.

- 1 Salzer, F. (1952/1982). *Structural hearing. Tonal coherence in music*. New York: Dover
2 Publications.
- 3 Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realization model.
4 *Cognition*, 58 (1), 75-125.
- 5 Schellenberg, E. G. (1997). Simplifying the implication-realization model of melodic
6 expectancy. *Music Perception*, 14 (3), 295-318.
- 7 Schellenberg, E. G., Adachi, M., Purdy, K. T., & McKinnon, M. C. (2002). Expectancy in
8 melody: Tests of children and adults. *Journal of Experimental Psychology*, 131 (4),
9 511-537.
- 10 Sergeant, D. (1983). The octave: Percept or concept. *Psychology of Music*, 11 (1), 3-18.
- 11 Stein D., & Spillman R. (1996). *Poetry into song: Performance and analysis of lieder*. New
12 York: Oxford University Press.
- 13 Temperley, D. (2007). *Music and probability*. Cambridge, MA: The MIT Press.
- 14 Temperley, D. (2008). A probabilistic model of melody perception. *Cognitive Science*, 32, 418-444.
- 15 Temperley D., & Marvin E. W. (2008). Pitch-class distribution and the identification of key.
16 *Music Perception*, 25 (3), 193–212.
- 17 Thompson, W. F., & Stainton, M. (1998). Expectancy in Bohemian folk song melodies:
18 Evaluation of implicative principles for implicative and closural intervals. *Music*
19 *Perception*, 15 (3), 231-252.
- 20 Thompson, W. F., Cuddy, L. L., & Plaus, C. (1997). Expectancies generated by melodic
21 intervals: Evaluation of principles of melodic implication in a melody-completion task.
22 *Perception & Psychophysics*, 59 (7), 1069-1076.
- 23 Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing
24 approach. *Psychological Review*, 107 (4), 885-913.

- 1 Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to
2 music: The dynamics of tonality induction. *Perception*, 32 (6), 741-766.
- 3 von Hippel, P. T. (2000). Redefining pitch proximity: Tessitura and mobility as constraints on
4 melodic intervals. *Music Perception*, 17 (3), 315-327.
- 5 von Hippel, P. T. (2002). Melodic-expectation rules as learned heuristics. In C. Stevens, D.
6 Burnham, E. Schubert, and J. Renwick (Eds.), *Proceedings of the 7th International*
7 *Conference on Music Perception and Cognition* (pp. 315–317). Sydney, Australia.
- 8 von Hippel, P. T., & Huron, D. (2000). Why do skips precede reversals? The effects of tessitura on
9 melodic structure. *Music Perception*, 18 (1), 59–85.
- 10 Zbikowski, L. M. (2002). *Conceptualizing music: Cognitive structure, theory, and analysis*. New
11 York: Oxford University Press.
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25

1 **Footnotes**

2 ¹ Krumhansl (1995) and Schellenberg (1997) suggest that Pitch Proximity should be
3 quantified in a graded fashion, as described here, but adding (instead of subtracting) one point
4 when distance in pitch between the current tone in a melody and the next tone increases by 1
5 semitone. If this coding method is used, the association between Pitch Proximity and any
6 measure of melodic expectation is expected to be negative (because higher values are given to
7 less expected tones). In order to avoid this discrepancy, in the present work Pitch Proximity
8 was quantified in a graded fashion, but with the opposite sign.

9 ² Schellenberg and his colleagues (Schellenberg, 1996, 1997; Schellenberg et al.,
10 2002)—and also Krumhansl (1995)—based their works on Narmour (1990)’s Implication-
11 Realization model (I-R) of melodic expectation; indeed, many of the terms taken from their
12 works (e.g., ‘pitch proximity’, ‘pitch reversal’, ‘registral direction’, etc. —see text) pertain to
13 this model. According to the I-R model, the expectations for tones proximate to the
14 penultimate tone heard become stronger as the distance in pitch between the last two tones
15 heard increases. That is to say, the model regards such expectations as being dependent not
16 only on the penultimate tone heard, but also on the last one. This would explain why
17 Schellenberg et al. (2002) propose that pitch proximity has a ‘second-order’ (instead of a
18 ‘lagged’) influence on melodic expectations.

19 ³ The tonal hierarchy index from Krumhansl and Kessler (1982) represents a
20 perceptual measure of the compatibility or stability of each pitch-class relative to each major
21 and minor key. It is based on probe-tone experiments in which 10 participants (out of which
22 only 1 had some formal music theory background) were played a key-establishing musical
23 context, such as a cadence or scale, and were asked to rate how well a probe-tone “fit” with
24 the context. Krumhansl and Kessler averaged the participants’ ratings across different
25 contexts and keys to create a single major key-profile and a single minor key-profile, which

1 jointly form the tonal hierarchy index used here. Figure 1 shows the major key-profile of the
2 tonal hierarchy index; the minor-key profile is 6.33, 2.68, 3.52, 5.38, 2.60, 3.53, 2.54, 4.75,
3 3.98, 2.69, 3.34, 3.17 (the highest value representing the tonic tone).

4 ⁴ The predictor variable ‘Tonal Hierarchy-6/5’ is referred to in this way because, based
5 on the previous melodic context, it codes tonal expectations in a range of a major or minor
6 ‘sixth’ (i.e., 9 or 8 semitones, respectively), or in a range of a perfect ‘fifth’ (i.e., 7 semitones).
7 In the example given in Figure 1, it codes tonal expectations in a range of a perfect fifth,
8 because the melody unfolds from F5 to Bb4 (as the potential anchors most proximate in
9 pitch). If the melody had unfolded, say, from D5 to F4 or from Bb4 to D4, the variable would
10 have coded expectations in a range of a major sixth or a minor sixth, respectively.

11 ⁵ Interestingly, in a series of experiments conducted by Krumhansl et al. (1987) it was
12 found that, when asked to rate how well probe-tones fitted with excerpts of atonal serial music
13 (in which, theoretically, the repetition of recent tones is avoided), listeners with more music
14 training rated the more recently sounded tones lower, whereas listeners with less music
15 training did the opposite. Furthermore, the more trained listeners gave lower ratings to probe-
16 tones consistent with the tonal implications of the musical excerpts (as measured by the tonal
17 hierarchy index from Krumhansl & Kessler, 1982), whereas the less trained listeners again did
18 the opposite. As noted by Krumhansl and her colleagues, it seems that the more trained
19 listeners implemented a strategy “of relating the serial contexts to tonal hierarchies of major
20 and minor keys, and simply reversing the ordering of the ratings” (pp. 73-74). Taken together,
21 these findings suggest that an interaction does exist between the conceptual knowledge
22 listeners have about music, the experimental task, and the musical materials used as stimuli.
23 This interaction might be the cause of the differences observed here between musicians and
24 non-musicians regarding their tonal expectations (see General Discussion).

1 ⁶ It is worth mentioning that in the analyses that Schellenberg (1996) made in his
2 Experiment 1 (in which listeners gave ‘continuation’ ratings for fragments ending not only on
3 small but also on large intervals), he found that tonality (oriented toward I/I; i.e., as coded in
4 Tonal Hierarchy) did not uniquely account for much of the variance in the data. Moreover, he
5 found that the influence of tonality was weaker on listeners with more music training than on
6 listeners with less music training (see Schellenberg 1996, Table 2), while previous studies
7 suggest that the opposite should be the case (e.g., Krumhansl & Shepard, 1979). As may be
8 noted, Schellenberg’s results also agree with the idea that tonality’s influence on expectation
9 is partially regulated by the conceptual framework listeners use to understand music.