PITCH: A KEY FACTOR IN TONALITY INDUCTION

J. FERNANDO ANTA
National University of La Plata, La Plata, Argentina

TWO EXPERIMENTS INVESTIGATED THE ROLE OF pitch-related information in tonality induction. In both experiments, participants were asked to: 1) identify (sing) the tonic of either an original sequence of tones or a distorted version in which pitch class distribution was preserved but pitch class ordering, pitch contour, and/or pitch proximity were altered; and 2) rate how confident they were in the tonic they identified. In Experiment 2, the sequences were presented with an isochronous rhythm, in order to eliminate the potential confounding effects of time-related information. The results of both experiments showed that participants' ability to identify the tonic of the sequences, as well as their confidence in the tonic they identified, decreased when pitch class ordering was distorted, and also when pitch proximity was reduced. This suggests that tonality induction not only involves the identification of abstract pitch class structures, but it also acts as a pattern-matching process.

Received: February 18, 2014, accepted July 11, 2014.

Key words: pitch class distribution, interval class vector, pitch class ordering, pitch contour, pitch proximity

TONALITY INFLUENCES MUSIC LISTENING IN profound ways. When listeners identify the key of a musical piece, they understand some tones and chords as being more or less stable than others. In particular, one tone, the *tonic*, is understood as the most stable pitch event, and provides a cognitive reference point around which the other pitch events are hierarchically organized (Krumhansl, 2004; Krumhansl & Cuddy, 2010; see however, Parncutt, 2011). This hierarchical organization, in turn, influences listeners' expectations; all things being equal, the more stable an event is, the more highly it is expected (Larson, 2004; Marmel, Tillmann, & Delbé, 2010; Pineau & Bigand, 1997; Schellenberg, 1996). Accordingly, the most stable events in a given key tend to be processed more quickly and accurately by listeners than the least stable ones (Bharucha & Stoeckig, 1986; Janata & Reisberg, 1988; Pineau & Bigand, 1997; Tillmann, Janata, Birk, & Bharucha, 2003). In addition, the encoding of tones and chords in a key affects musical memory and recognition (Bharucha & Krumhansl, 1983; Krumhansl, 1979; Krumhansl & Castellano, 1983; Schmuckler, 1997), as well as listeners' judgments about the emotional content of music (Costa, Fine, & Ricci Bitti, 2004; Gabrielsson & Lindström, 2010; Sloboda, 1991).

Given the influence that tonality has on music listening, a major issue in music psychology relates to how listeners identify the key of a musical piece. The question is: how does tonality induction work? This question has mainly been addressed here by focusing on how listeners identify the tonic in Western tonal music. A number of ideas have been proposed about this issue. However, two of them, usually referred to as the *distributional* and *functional* views (e.g., Brown, 1988; Temperley & Marvin, 2008), have been particularly influential, stimulating a wealth of theoretical and empirical research. These views also stimulated the present work, whose overall aim was to shed light on tonality induction by clarifying their hypotheses and overcoming their limitations.

The Distributional and Functional Views of Tonality Induction: Theory, Evidence, Criticisms

According to the distributional view, tonality induction depends on the distribution of pitch classes in music. Perhaps its most popular version is the Krumhansl-Schmuckler key-finding algorithm (Krumhansl, 1990). This algorithm is based on three hypotheses. The first is that listeners are sensitive to the total duration accumulated by pitch classes in music. The second is that, in psychological terms, keys are cognitive templates or *key profiles* that represent the ideal (or average) duration-based distribution of pitch classes in tonal pieces–in such a way that the longer the duration accumulated by a given pitch class is, the higher is its level of tonal stability. Two key profiles are proposed, one for major keys and one for minor keys; Figure 1 (top panels, filled circles) shows these profiles. Finally, the third hypothesis is that to find the key of a musical piece, listeners compare (although not necessarily consciously) the duration-based distribution of pitch classes in the piece with the key profiles shown in Figure 1 and choose the

Music Perception, VOLUME 32, ISSUE 4, PP. 413–433, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312. DOI: 10.1525/MP.2015.32.4.413

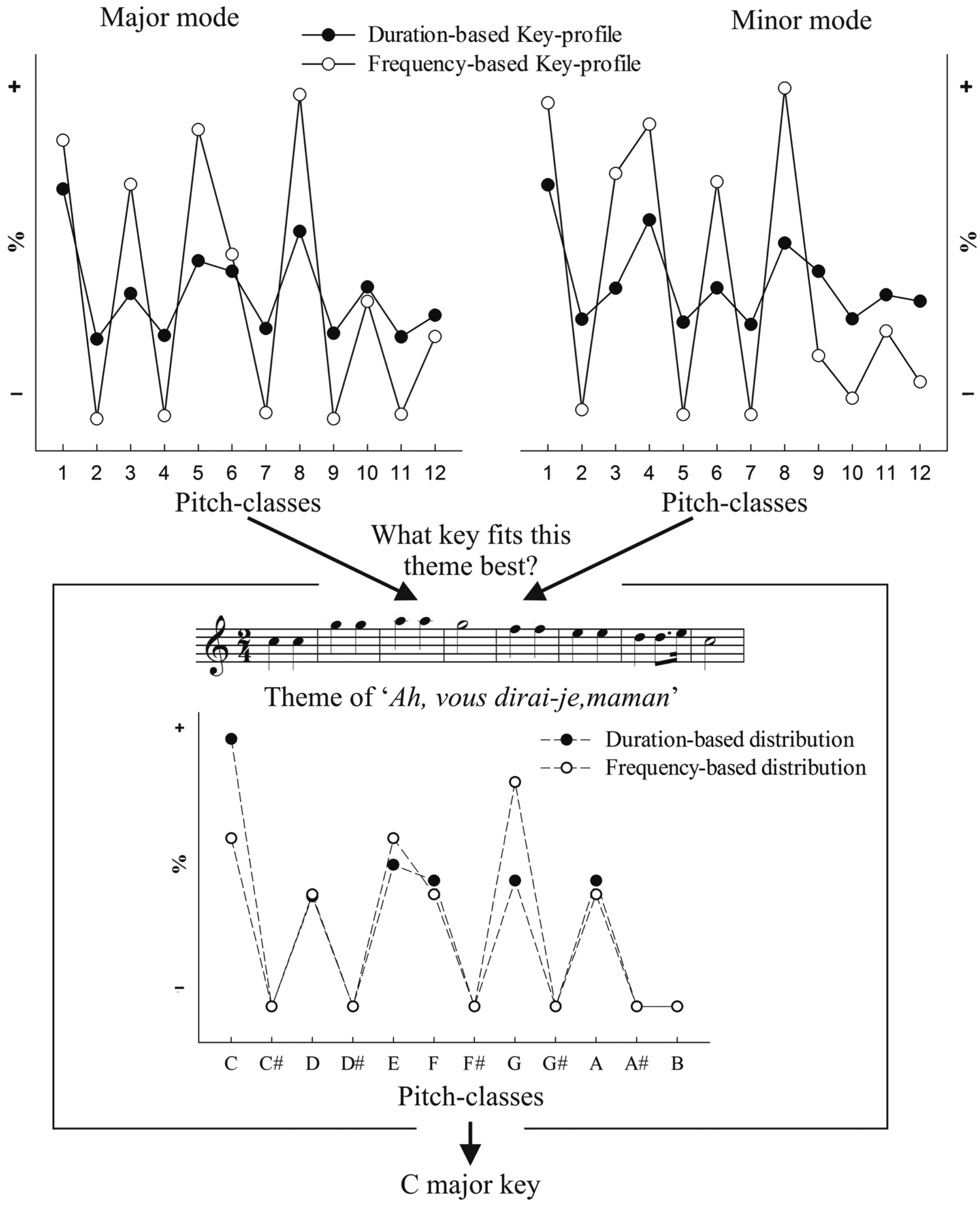

FIGURE 1. Schematic representation of the process of tonality induction as suggested by the duration-based (filled circles) and the frequency-based (open circles) versions of the distributional view. In both versions, key profiles representing the distribution of pitch classes in tonal music (top panels) are compared with the distribution of pitch classes in the current musical piece (bottom panel), in this case, the theme of *Ah, vous dirai-je, maman* (by W. A. Mozart). Both versions predict that listeners will perceive the theme in the key of C Major.

key (profile) that yields the highest correlation coefficient. For instance, Figure 1 (bottom panel, filled circles) also shows the theme of *Ah, vous dirai-je, maman* and its duration-based pitch class distribution: the algorithm predicts that listeners will perceive the theme in C Major, because the highest correlation is achieved when the pitch class distribution of the theme is compared with a major key profile in which the tonic is C.

There is another version of the distributional view that has been successfully developed. It is similar to Krumhansl-Schmuckler's version, except that it claims that tonality induction depends on the frequency of occurrence (or onset, instead of the total duration) of pitch classes (e.g., Bharucha, 1987; Temperley, 2008; Tillmann, Bharucha, & Bigand, 2000). Although the superiority of one version over the other is currently a matter of debate (Smith & Schmuckler, 2004; Temperley & Marvin, 2008; see also Krumhansl & Cuddy, 2010), both have been able to account for listeners' responses in key-finding tasks. In fact, they usually lead to the same predictions. For instance, Figure 1 also shows the frequency-based key profiles generated from the *Essen Folksong Collection* (Schaffrath, 1995) by Temperley (2008), and the frequency of occurrence of pitch classes in the theme (top and bottom panels, respectively, open circles). Again, it is predicted that listeners will perceive the theme in C Major, because the highest correlation is achieved when its (frequency-based) pitch class distribution is compared with a major (frequency-based) key profile in which the tonic is C.

In spite of its success, the distributional view, as a whole, has been sharply criticized, mainly because it disregards the fact that intervals between pitches may provide listeners with important clues to find the key. This criticism is the hallmark of the functional view. According to this view, tonality induction depends on interval class patterns and pitch class ordering that pitches form. Initially, Browne (1981) suggested that listeners are sensitive to the interval classes present in a musical piece and that, in psychological terms, keys are cognitive templates that represent the interval class vectors that characterize tonal pieces. For example, major keys would represent the vector 2-5-4-3-6-1, which means that in major-key pieces, intervals may be reduced to 2 minor seconds, 5 major seconds, etc.–a diatonic set is assumed. The important point here is that the various interval classes occur a different number of times in the vector and that, consequently, the rarer (or less frequent) interval classes only involve a few pitch classes. In particular, only pitch classes a perfect fourth above and a minor second below the tonic tone can form the tritone, the rarest interval class in the vector. Therefore, Browne (1981) concludes, the rarest interval classes improve tonality induction by providing the strongest cues to the tonic.

Evidence supporting Browne (1981)'s hypothesis was reported by Brown and Butler (Brown, 1987, 1988; Brown & Butler, 1981), who found that listeners' ability to identify the tonic increases when a tone sequence presents one tritone, as opposed to none. Moreover, they found that the ability decreases when the sequence presents more than one tritone, suggesting that, given its rarity, each tritone implies a different tonic. However, Butler and Brown also observed that, in identifying the key, listeners are sensitive to the temporal ordering of tones. Specifically, listeners' ability to identify the tonic increases when the tones of the tritone occur in the order subdominant/leading tone, or when one of them occurs at the end of the sequence. Hence, they argued for a functional view in which both interval class patterns and pitch class ordering are crucial (see also Brown, Butler, & Jones, 1994; Matsunaga & Abe, 2005). This view, however, may also be criticized. For instance, it has some difficulties in finding the key of the theme shown in Figure 1, where there is no tritone. Furthermore, this fact strongly contrasts with the tonal clarity of the theme, whose tonic appears to be obviously C.

In the past few years, important efforts have been made to combine the distributional and functional views of tonality induction to overcome their limitations. However, the success of these efforts has been mixed. Huron and Parncutt (1993), for example, sought to improve the Krumhansl-Schmuckler algorithm by taking into account subsidiary pitches (harmonics) of tones and the effects of temporal ordering–via the influence of echoic memory for pitch. Interestingly, the ensuing algorithm performed better than its original counterpart in predicting key implications of harmonic sequences, but it was still unable to account for different key implications arising from reordering pitch classes in tone sequences. Tillmann et al. (2000) designed several frequency-based algorithms in which subsidiary pitches and pitch class ordering were considered either separately or jointly. They found that algorithms' outcomes did not vary qualitatively. Similarly, Toiviainen and Krumhansl (2003) compared an algorithm based on pitch class distribution and another based on pitch class ordering with a third algorithm in which the former two were combined, and observed that the combined algorithm only slightly (about 1%) improved the performance of either single algorithm alone.

What is the missing link between the distributional and functional views? What factor should be incorporated to understand tonality induction better? However paradoxical it may sound, the factor proposed here is pitch.

Indeed, although pitch may clearly be viewed as a key factor for tonality induction, both views posit that the whole process relies on information about pitch classes–and not about pitches. This is so because, basically, they both assume that tones separated by octaves are perceptually (and then tonally) equivalent. Thus, for example, from the distributional viewpoint what matters is how long or how often the pitch class C occurs in the theme shown in Figure 1, and not whether the tone that actually occurs is C4, C5, or any other tone of the class C. Similarly, from the functional view what matters is whether the pitch classes C and B are present in the theme, and not whether the interval between them is a minor second, a major seventh, or any other interval of the class "1"–i.e., "minor second." It is well established that octaves are particularly consonant or "smooth" (e.g., Schellenberg & Trehub, 1994; see also Shepard, 1964–c.f. Burns, 1981), and that octave relations provide several processing advantages–e.g., in tune recognition (Francès, 1958), or in mistuning identification (Lee & Green, 1994; see also Deutsch, 2013). However, there is evidence that only musically trained listeners consistently judge octaves as highly similar relative to other pairs of tones (Allen, 1967; Sergeant, 1983), and also that their processing advantages are not fully operative in untrained listeners (Burns & Houtsma, 1999; Elliot, Platt, & Racine, 1987). Further, there is evidence that, regardless of music training, octave equivalence effects may not operate directly in the processing of melodic sequences (Deutsch, 1972; Deutsch & Boulanger, 1984; Kallman, 1982; see also Deutsch, 1969). Thus, it seems reasonable to consider the possibility that tonality induction might not depend entirely on pitch class related information and, therefore, that pitch-related information might support this process. Two experiments were conducted to investigate this possibility.

Experiment 1

At this point, what is meant by "pitch-related information" must be clarified. If two tone sequences differ only in one pitch, then the pitch-related information they communicate may differ with regard to two factors: pitch contour and pitch proximity. Basically, pitch contour refers to the pattern of ups and downs from one tone to the next, whereas pitch proximity refers to the distance (in pitch) between them.[1] There is abundant evidence demonstrating that the distortion of pitch contour and pitch proximity affects musical processing (for a recent review, see Halpern & Bartlett, 2010). More importantly, there is evidence that these kinds of pitch-related information interact with pitch class related information in determining what is understood as 'tonal' music. Most of this evidence comes from theoretical and pedagogical treatises on Western tonal harmony, but there is some psychological and behavioral evidence as well.

For example, musical treatises (e.g., Aldwell & Schachter, 2003; Piston, 1941/1959) state that in order for music to sound tonal, dissonances must be *resolved* (i.e., followed by consonances). This is thought to be particularly important in the case of the tritone, the most dissonant interval. As Piston (1941/1959) put it, tritones are to be almost invariably resolved either by the leading tone *continuing one step upward* to the tonic tone, or by the subdominant tone *continuing one step downward* to the mediant tone. Similarly, it is argued that melodies are to be kept within a *small range* and that, whenever possible, *skips* are to be avoided. Furthermore, it is also argued that when a melody moves by skip, the best procedure is to change pitch contour immediately after, or alternatively to use the same harmony for both tones. Clearly, these "rules of voice leading," as musicians call them, imply that in order for a tonal center to be unambiguously established, pitch contour and pitch proximity must be handled in specific ways. In fact, it has been suggested that a few pitch-related factors, in particular pitch proximity, largely explain the rules of voice leading upon which tonal music is built (Huron, 2001), which implies that, to some extent, they might also account for the "tonicization" of tonal music.

In line with these ideas, Cuddy and colleagues (Cuddy & Lyons,1981; Cuddy, Cohen, & Mewhort, 1981) found that even a simple tonal structure (e.g., I-V-I) is difficult to recognize as a perceptual unit if a tone sequence does not fit well with common rules of voice leading and tonal patterning. Interestingly, these rules often reflect important pitch-related expectations the listener generates when hearing a tonal context–e.g., expectations for an upward semitone after the seventh scale degree, or for a reversal in pitch direction after a skip (Larson, 2004; Schellenberg, 1996)–which may explain why, once such rules are broken, the sense of tonality become weaker. In addition, Boltz and Jones (1986) and Rosner and Meyer (1986) found that listeners' ability to encode a tonal sequence depends on some features of pitch contour, especially on its peaks; seemingly, contour peaks serve as accents that delineate the (tonal) goals of musical flow (see also Lerdahl & Jackendoff, 1983).

On the other hand, Bharucha (1984; see also Deutsch, 1984) showed that listeners' judgments about the musical

---

[1] In some recent studies, pitch contour and pitch proximity have been described more globally, being referred to pitch relations between *nonadjacent* tones in time and/or between more than two tones (e.g., Quinn, 1999; von Hippel, 2000). Further, they (plus other factors) have been mixed to compound a unified description of melodic contour (e.g., Schmuckler, 1999, 2010). The choice of focusing on local (and separate) descriptions of pitch contour and pitch proximity was primarily pragmatic: they are more easily manageable, thus providing a suitable starting point for examining the role of direction and distance between pitches in tonality induction.

fit between tones and chords tend to reflect tonal conventions only when two conditions are met: first, a progression from a nonchord tone to a chord tone (and not in the reverse order); second, the tonally stable tone is proximate in pitch to the tonally unstable tone. As may be noted, the first condition lends further support to the functional view of Brown and Butler (1981). However, the second condition suggests that pitch proximity also affects the sense of tonality. In line with this, in our laboratory we have recently found that listeners' melodic expectations consistently reflect the influence of tonality only when (upcoming) tones are proximate in pitch (Anta, 2013), which suggests that, to some extent, pitch proximity consolidates that influence.

To summarize, the distributional and functional views of tonality induction posit that the key-finding process depends on pitch class-related information. However, there is evidence suggesting that this process does not depend entirely on pitch classes. Pitch-related information then may be thought of as a natural candidate for supporting key finding. Indeed, there is evidence that information about pitch contour and pitch proximity contributes to the tonicization of music. This leads one to hypothesize that pitch contour and pitch proximity might facilitate or inhibit tonality induction. In order to test these hypotheses, the following experiment was conducted.

METHOD

Participants. Thirty five undergraduate students of music from the Faculty of Fine Arts of the National University of La Plata (UNLP), Argentina, took part in the experiment voluntarily. The participants' mean age was 25 years (range: 20-32 years). They had a mean of 7 years of formal training (range: 3-15 years), including ear training, music theory, and music performance. They also had informal experience singing or playing an instrument, with a mean of 10 years (range: 4-20 years). None of the participants reported having absolute pitch or vocal impediments to singing.

Stimuli and apparatus. Thirty-six tone sequences were used as stimuli. Six sequences were taken from the vocal lines of *Lieder* by Franz P. Schubert, which may be seen as representative of the Western tonal music of the common-practice period (Piston, 1941/1959); they will be referred to hereafter as the *Original* sequences (see Figure 2). Half of the *Original* sequences were in a major key (G, E, and Eb major) and half in a minor key (F#, D, and C minor). These sequences were selected, first, because they were "tonally efficient," that is, because they would tend to convey the intended key (i.e., the key indicated by the key signature). Tonal efficiency was determined with standard analytical procedures (e.g., based on key signature), but it was supported by the procedures proposed by the distributional and functional views. The Krumhansl-Schmuckler algorithm yielded the highest correlation for the intended key in five of the six sequences (mean $r = .77$, ps $< .05$); the exception was the sequence in E major. Similarly, a frequency-based algorithm based on the key profiles generated by Temperley (2008) yielded the highest correlation for the intended key in four of the six sequences (mean $r = .81$, ps $< .05$); the exceptions were the sequences in E major and C minor. However, all the sequences had the tritone of the intended key; indeed, the major key sequences only had that tritone.

Additionally, the *Original* sequences were selected because they met three other criteria. First, they remained in the same key from the beginning to the end (i.e., they had no "chromatic" pitches, according to key signature–except for the tone G5 in the F# minor sequence, which is quickly subsumed into the descent form of the "melodic" minor scale). Thus, the potential confounding effect of modulation on tonality induction was eliminated. Second, they had a similar duration–i.e., a roughly similar number of events (11, 13, 14, 15, 16, or 20 events) distributed over a roughly similar number of beats (9, 7, 13, 6.5, 6.5, and 9 beats, respectively). Thus, the potential confounding effect of sequence length was reduced. Finally, they started at the beginning but ended before the completion of a melodic group, on a tonally unstable tone–the second, fourth, or seventh scale degree. To meet this criterion, the last tone of Schubert's musical phrases, which was either the tonic or the mediant tone, was not included in the sequences: otherwise, it would have been easy for listeners to find their tonics–which was part of the experimental task (see *Procedure*). Moreover, given that when perceiving a tone sequence listeners show some preference for the tones heard most recently (Greene & Samuel, 1986; Surprenant, 2001; see also Butler, 1989), by removing the last tone of Schubert's musical phrases another potential confounding variable was eliminated: if participants found the tonic of the sequences, that could not be attributed to the fact that the tonic (or any other member of the tonic chord) was the last tone they heard.

The remaining thirty sequences used as stimuli were distorted versions of the *Original* sequences. Specifically, from each *Original* sequence five sequences were generated (see Figure 3). In the *Unordered* sequences, the order that pitches had in the *Original* sequences was randomly altered. In the *Original-dispersed* and *Unordered-dispersed* sequences, the pitches of the
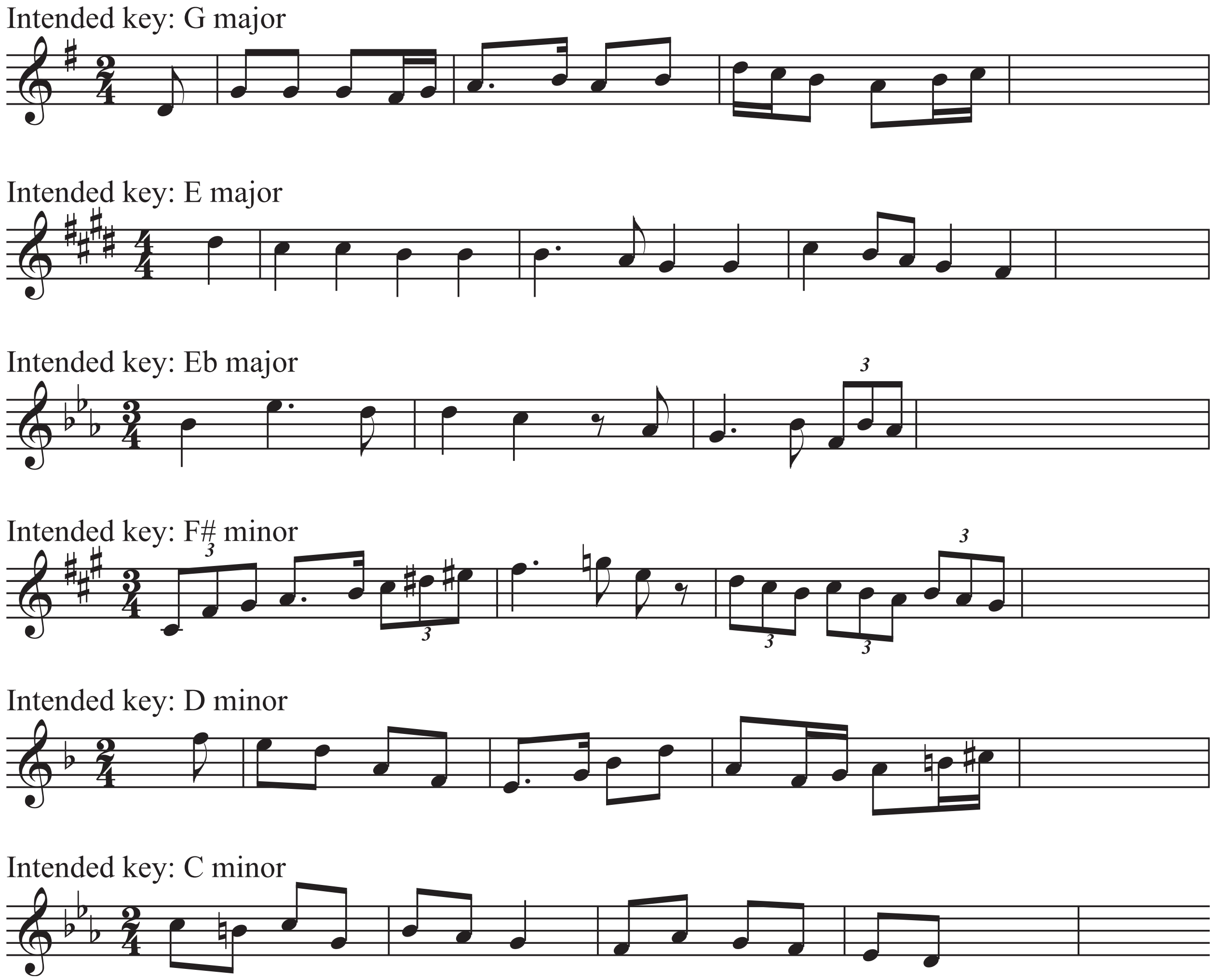
FIGURE 2. The six original sequences used in Experiment 1. The sequences were extracted from the vocal line of *Lieder* by Franz P. Schubert, Opus 89 (Breitkopf & Härtel, Leipzig; serie 20, N° 517-540): numbers I, IV, XVI, VI, I, and XV, respectively. (In the original score, the key signature of the sequence in E major was Eb major: the sequence was transposed to E major here to avoid that two original sequences have the same tonic).

Original and *Unordered* sequences, respectively, were randomly transposed by octaves within a two-octave pitch range. Finally, in the *Original-expanded* and *Unordered-expanded* sequences, the pitches of the *Original* and *Unordered* sequences, respectively, were randomly transposed by octaves within a 3-or-more octave pitch range in order to preserve pitch contour. Thus, the *Original* and *Unordered* sequences differed in pitch class ordering and pitch contour, but were similar in pitch proximity (mean interval between adjacent pitches, measured in semitones: 2.10 and 2.52, respectively). The *Original* and *Original-dispersed* sequences as well as the *Unordered* and *Unordered-dispersed* sequences differed in pitch contour and pitch proximity (mean: 2.10 versus 10.53, and 2.52 versus 10.25, respectively), but not in pitch class ordering. The *Original* and *Original-expanded* sequences as well as the *Unordered* and *Unordered-expanded* sequences differed in pitch proximity (mean: 2.10 versus 13.23, and 2.52 versus 12.37, respectively), but not in pitch class ordering or pitch contour. Finally, the *Original-dispersed* and *Original-expanded* sequences as well as the *Unordered-dispersed* and *Unordered-expanded* sequences differed in pitch contour, but not in pitch class ordering, and only slightly in pitch proximity (mean: 10.53 versus 13.23, and 10.25 versus 12.37, respectively). (It was assumed that these relatively small differences in pitch proximity, as well as those between the *Original* and *Unordered* sequences–less than one semitone–would not affect participants' responses. Indeed, the results supported this assumption–see below.)
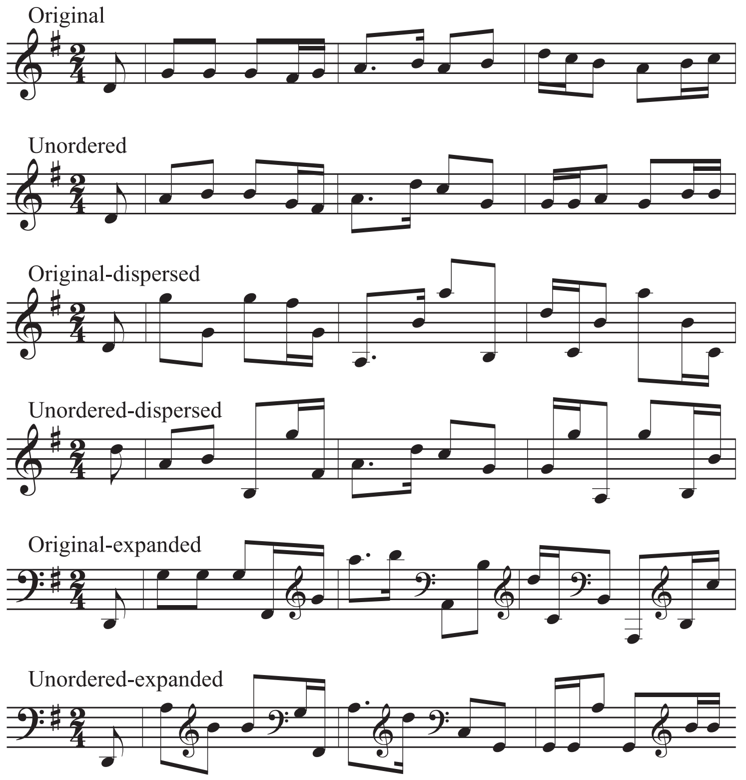
FIGURE 3. One of the six sets of melodic sequences used in Experiment 1. Each set contained one original sequence and five distorted versions of it. (Distorted sequences were randomly transposed to different keys before each experimental session; see text).

It is important to note, therefore, that owing to the same/different relationships involved in the materials, pitch proximity's effect on tonality induction, if existent, would be easily observable: participants' performance should be worse when hearing the *Original-expanded* sequences than when hearing the *Original* ones. However, pitch contour's effect, if existent, would be more difficult to observe, since there was no distorted sequence in which, compared to the corresponding original sequence, only pitch contour were distorted–because when pitch contour was distorted, pitch proximity was distorted too. Thus, there were two possible scenarios by which the effect of pitch contour could be demonstrated. In the first scenario, participants' performance should remain roughly constant when hearing the *Original* and *Original-expanded* sequences, but it should vary when hearing the *Original-dispersed* ones: this would imply that pitch proximity has no effect on tonality induction and, therefore, that the difference observed with the *Original-dispersed* sequences was due to pitch contour. In the second scenario, participants' performance should vary when hearing the *Original* and *Original-expanded* sequences, but even more when hearing the *Original-dispersed* ones: this would imply that the effects of pitch proximity and pitch contour on tonality induction were somehow additive.

It is important to note also that for generating the distorted sequences, the rhythmic structure of the *Original* sequences was held constant. Therefore, according to the distributional view any distorted sequence would convey the intended key (or any other key) with the same strength as its *Original* counterpart, because the duration-based and frequency-based distributions of pitch classes were the same.[2] According to the functional

---

[2] Strictly speaking, in the three major-key melodic sets the frequency-based distribution of pitch classes was slightly modified when pitch class ordering was modified, because in each case one single note–e.g., a quaver (or an eighth note)–was divided into two notes–e.g., two semiquavers (or sixteenth notes)–and vice versa. For example, in the *Original* and *Unordered* sequences shown in Figure 2 the pitch class G occurs four and five times, respectively, because one quaver was divided into two semiquavers, and the pitch class C occurs 2 times and 1 time, respectively, because two semiquavers were fused into one quaver. Notwithstanding this, the correlations between the frequency of occurrence of pitch classes in the sequences and the corresponding frequency-based major key profile were still significant ($ns = 12, ps < .05$) and, on average, remained roughly constant (mean $r = .76$ and .74, for *Original* and *Unordered* sequences, respectively).

view, any sequence in which pitch class ordering was distorted would be tonally more ambiguous. However, the *Original-dispersed* and *Original-expanded* sequences as well as the *Unordered-dispersed* and *Unordered-expanded* sequences would convey the intended key with the same strength as their *Original* and *Unordered* counterparts, because they had the same interval class vector and pitch class ordering.

Next, once distorted sequences were generated, each of them was transposed to a distant (or perceptually unrelated) key from the intended key of its *Original* counterpart, in order to reduce the potential confounding effect of tonal relatedness. Indeed, it has been shown that a given key is perceived as more closely related with its relative key or with the neighboring keys on the circle of fifths, than with the remaining keys (Krumhansl & Kessler, 1982). Therefore, these closely related keys were avoided and other, less related keys were selected to obtain the distorted sequences. Specifically, distorted sequences were randomly transposed by one of the following (upward) intervals: minor second, augmented fourth, major sixth, minor seventh, and major seventh, when the key of the original sequence was major; and minor second, major third, augmented fourth, minor seventh, and major seventh, when the key of the original sequence was minor.

Finally, stimuli were coded as MIDI files using Finale® 2010, generated as sound tracks with Garritan Instruments for Finale® set to the “Steinway Piano” timbre, and recorded as MP3 files. Each *Original* sequence was adjusted to the tempo suggested in the score from which it was extracted, and the same tempo was assigned to its distorted versions. All the sequences were presented through external speakers (JVC/MIX-J10 speakers), at a dynamic level adjusted according to participants’ preference.

Procedure. In individual sessions, participants were told that they would hear several melodic fragments, and that after hearing each fragment they would be asked to perform two tasks: first, to sing the tonic of the fragment; and second, to rate on a scale from 1 (*little confident*) to 7 (*very confident*) how confident they were that the tone they had sung was actually the tonic. After receiving the instructions, the participants heard three practice trials; these trials were randomly selected from original and distorted sequences based on three other melodies, equivalent to those described in the “Stimuli and Apparatus” section. Next, participants were asked if they had understood the tasks and if they felt that they could perform them well; all of the participants answered these questions affirmatively. Finally, they heard thirty-six test trials, one for each of the thirty-six tone sequences described above.

Each trial began with a white-noise of 2 s, to reduce the influence of memory (for the key of the last sequence heard) on subsequent response; the white-noise was followed by a silence of 2 s and, finally, by a tone sequence. After the sequence was heard, each participant sang her/his response. (To reduce the likelihood that responses were inhibited by performance limitations, participants were allowed to vocalize as many pitches as needed to find the tonic, and to support their responses with verbal descriptions). Then, the experimenter (who was blind as to the true tonic of each trial) searched for the pitch the participant had sung (or dictated via verbal description) on the virtual keyboard of Garritan® (set to the Steinway Piano timbre used in generating the sequences); the search ended when the participant confirmed that the pitch she/he had identified as the “tonic” was the pitch the experimenter played on the virtual keyboard. Next, the experimenter listed on a paper sheet the tonic the participant had identified, and asked her/him how confident she/he was in the response. Finally, the experimenter listed the level of confidence and, once the participant confirmed that she/he was ready, a new trial began.[3]

Within each group of trials, for practice or testing, the order of trials was random and different for each participant. However, two exceptions were made in the randomization process: tone sequences belonging to the same set (e.g., the *Original* and *Unordered* sequences shown in Figure 3) or having the same intended tonic (e.g., G, if the intended key was G major or G minor) were never presented immediately after each other. Each session was carried out in a sound attenuated room, where participants sat in front of a blank wall, and lasted approximately 45 min. At the end of the session, participants filled out a questionnaire about their musical background. (It is worth mentioning that informal post-test interviews suggested that the participants were not familiar with the original melodies used as stimuli;

---

[3] At this point, it has to be mentioned that the procedure used here was largely borrowed from Brown (1988). Basically, the only difference was that participants’ responses were not recorded as audio files for future verification, as done in Brown’s work, but verified *in situ* by participants themselves. Responses verification was done this way because preliminary tests (in which responses were recorded) showed that, occasionally, participants’ responses departed slightly from the tempered tuning system: hence participants were required to verify themselves their responses and then to interpret themselves the departures. It is worth mentioning that the data collected in the (fifteen) preliminary tests yielded a pattern of results that are consistent with the main findings reported below.

indeed, none of them indicated that they recognized the melodies as extracted from Schubert's songs, even when many of them were specifically asked about the issue).

RESULTS AND DISCUSSION

Only responses from test trials were analyzed herein. The first question addressed was which tones were preferred (i.e., selected most often) by the participants as tonic tones when they were given the *Original* sequences. Out of the 35 participants, 29 preferred the intended tonics (selecting them in four or more sequences). On average, intended tonics were selected 78.6% of the times. This indicated that the *Original* sequences tended to convey the tonics they were expected to convey. Thus, in subsequent analyses the intended tonics for the *Original* sequences were considered as Correct Tonics (CT), whereas other tonics were considered as Incorrect Tonics (IT).

Next, the main question of the experiment was addressed, namely whether participants' preference for CT varied significantly when they were given the distorted sequences instead of the *Original*. To address this question, average proportions of CT (expressed as percentages) from each participant were entered into a one-way repeated-measures analysis of variance (ANOVA) with Melodic Condition (six levels: *Original*, *Original-dispersed*, *Original-expanded*, *Unordered*, *Unordered-dispersed*, and *Unordered-expanded*) as the within-subjects factor. Figure 4 (upper panel) depicts the results of these analyses. The effect of Melodic Condition was significant, $F(5, 30) = 38.41$, $p < .001$. Bonferroni post hoc comparisons showed that the proportions of CT were higher in the *Original* condition than in the remaining conditions (all ps < .001). This means that the alterations introduced in the *Original* sequences reduced participants' ability to identify the intended tonics–and, therefore, the intended keys.

To ensure that this result was not an artifact of averaging data across melodic sets, data from each set (a melody and its altered versions) were entered separately into Friedman's one-way repeated-measures ANOVAs with Melodic Condition as the within-subjects factor. The results of these analyses are summarized in Table 1 (upper portion). As may be seen, in all the sets the proportions of CT in the *Original* condition were higher than in the remaining conditions. This confirmed that participants' ability to find the key decreased when the sequences were distorted. (Differences between sets were explored further by testing the effects of some additional factors–e.g., the tonal strength of each set, indexed by the Krumhansl-Schmuckler algorithm–but results were not significant.)

Two other results depicted in Figure 4 (upper panel) are particularly important. First, the proportions of CT in the *Original-dispersed, Original-expanded*, and *Unordered* conditions did not differ significantly. The relevance of this result is twofold. On the one hand, it suggests that the distortion of pitch contour did not affect tonality induction: otherwise, participants' preference for CT should have been lower in the *Original-dispersed* condition, where both pitch contour and pitch proximity were distorted, than in the *Original-expanded* condition, where pitch proximity was distorted but pitch contour was preserved. On the other hand, it suggests that the distortion of pitch proximity and pitch class ordering inhibited tonality induction to a similar degree. Second, the lowest proportions of CT were found in the *Unordered-dispersed* and *Unordered-expanded* conditions, suggesting that the effects of pitch class ordering and pitch proximity were additive. Indeed, a two-way repeated-measures ANOVA with Order (two levels: *original*, *random*) and Transposition (three levels: *none*, *2-octaves, 3-or-more-octaves*) as within-subjects factors showed that both main effects were significant, $F(1, 34) = 136.01$, and $F(2, 33) = 32.11$, respectively; ps < .05; the proportion of CT were higher when the sequences had the original orderings, or the shortest distances (i.e., none transposition) between successive tones. However, their interaction was not significant, $F(2, 33) = 0.84$, $p = .44$.

In subsequent analyses, Incorrect Tonics (IT) were examined. Regarding these analyses, one has to bear in mind that, as mentioned above (see "Stimuli and Apparatus" section), a given key is perceived as more closely related with its relative key or with the neighboring keys on the circle of fifths, than with the remaining keys. In addition, it should be pointed out that the most stable tones in a key (i.e., the components of the tonic chord) are perceived as more closely related to each other than to the remaining tones (Krumhansl, 1979, 1990). Therefore, one would expect the participants to confound, for example, the (correct) tonic C with the "tonally related" IT, A, G, F, or E, more often than with the remaining, "tonally unrelated" IT, C#, D, D#, F#, G#, A#, and B (C major being the intended key). Further, given the differences observed between conditions one would expect the participants to confound CT with "tonally related" IT more often than with "tonally unrelated" IT when they were given the *Original* instead of the distorted sequences. To assess these possibilities, IT were entered into a two-way repeated-measures ANOVA with Tonal Status (two levels: *related*, *unrelated*) and Melodic Condition as within-subject factors. It is worth noting that Tonal Status was treated as an

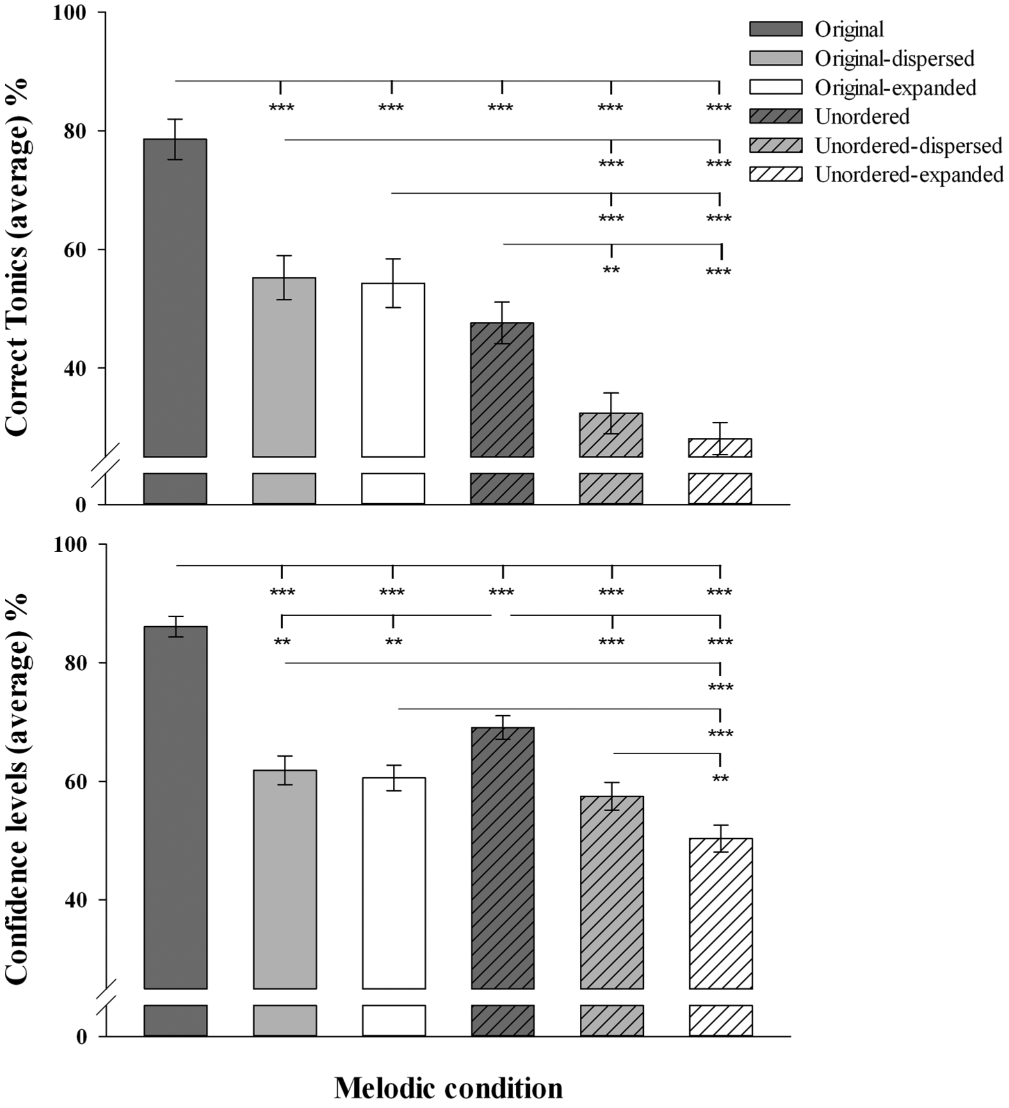

FIGURE 4. Average proportions (in %) of Correct Tonics and Confidence Levels observed in each melodic condition in Experiment 1. Asterisks indicate statistical significance of the difference between conditions (** $p < .01$, *** $p < .001$). Error bars represent SEMs.

independent variable: the dependent variable was the number of trials on which each type of IT was chosen. (There were six test trials in which four different participants did not respond; i.e., did not sing a tonic. These cases were considered as "unrelated" incorrect responses).

As predicted, there was a significant main effect of Tonal Status, $F(1, 34) = 63.20$, $p < .001$; overall, related IT were selected more often than unrelated IT (68.9% and 31.1% of the times, respectively). The main effect of Melodic Condition was also significant, $F(5, 30) = 44.59$, $p < .001$, with fewer errors in the *Original* condition (7.0%) than in the other conditions (mean 18.6%; range 14.7%–23.9%). There was a significant interaction between Tonal Status and Melodic Condition, $F(5, 30) = 2.55$, $p = .049$. However, this interaction was not as predicted. The analysis of simple main effects revealed that differences between related and unrelated IT were larger in the distorted conditions (mean 6.8%; range 5.6%–9.7%) than in the *Original* condition (3.9%), suggesting that the sense of tonal relatedness between keys was stronger among distorted sequences. This result, however, should be interpreted in the context of the relatively small proportion of errors observed in the *Original* condition, as compared to those observed in the remaining conditions. (This point will be discussed again in Experiment 2).

TABLE 1. Average proportions (in %) of *Correct Tonics* and *Confidence Levels* observed in each Melodic Set and each Melodic Condition in Experiment 1.

Correct Tonics

	Melodic Condition					
Melodic Set	Original	Original-dispersed	Original-expanded	Unordered	Unordered-dispersed	Unordered-expanded
G major*	85.71	65.71	80.00	60.00	51.43*	48.57*
E major**	57.14	40.00	34.29	42.86	8.57***	22.86
Eb major***	85.71	65.71	34.29***	74.29	54.29*	54.29*
F# minor***	82.86	40.00**	51.43*	22.86***	14.29***	2.86***
D minor***	88.57	54.29*	60.00*	40.00**	28.57***	11.43***
C minor*	71.43	65.71	65.71	45.71	37.14*	28.57**
Mean	78.57	55.24***	54.29***	47.62***	32.38***	28.10***
SEMs	6.69	8.29	8.04	8.09	7.33	6.73

Confidence Levels

	Melodic Condition					
Melodic Set	Original	Original-dispersed	Original-expanded	Unordered	Unordered-dispersed	Unordered-expanded
G major***	92.65	58.78***	67.35***	79.18***	60.41***	59.18***
E major***	85.31	63.27***	60.82***	68.57*	55.10***	50.20***
Eb major***	82.86	64.90**	50.20***	77.96	64.90***	53.06***
F# minor***	81.63	57.96***	57.96***	55.51***	49.80***	44.90***
D minor***	84.90	57.96***	57.96***	64.90***	54.29***	46.12***
C minor***	88.98	68.16***	68.98***	68.16***	60.41***	48.57***
Mean	86.05	61.84***	60.54***	69.05***	57.48***	50.34***
SEMs	2.64	3.68	3.55	3.10	3.48	3.63

* $p < .05$, ** $p < .01$, *** $p < .001$.

Note: For each Melodic Set, asterisks indicate statistical significance of the effect of Melodic Condition. For each Melodic Condition, asterisks indicate statistical significance of the difference between values in the Original condition and the marked condition; values in unmarked conditions did not differ significantly from those observed in the corresponding Original condition.

Finally, participants' confidence ratings were analyzed. Specifically, the confidence ratings reported by each participant for each selected tonic were averaged within each melodic condition, and average confidence ratings were entered into a one-way repeated-measures ANOVA with MELODIC CONDITON as the within-subjects factor. Figure 4 (lower panel) depicts the results of this analysis. In line with the results from the tonic identification task, there was a significant effect of Melodic Condition, $F(5, 30) = 38.92$, $p < .001$. Bonferroni post hoc comparisons showed that the levels of confidence reported in the *Original* condition were higher than those reported in the other conditions (all ps $< .001$). To further examine this finding, confidence ratings from each melodic set were analyzed separately using one-way repeated-measures ANOVAs. Table 1 (lower portion) summarizes the results of these analyses. As shown in the table, in all the sets the levels of confidence were higher in the *Original* condition than in the remaining conditions.

Figure 4 (lower panel) shows three other results from the confidence level task that deserve special attention. First, the levels of confidence in the *Original-dispersed* and *Original-Expanded* conditions did not differ significantly. This supports the idea that the distortion of pitch contour in the *Original-dispersed* condition did not affect tonality induction. (It is not clear, however, why the difference between confidence levels in the *Unordered-dispersed* and *Unordered-expanded* conditions was significant. This difference might reflect that, overall, tones were closer to each other in pitch in the *Unordered-dispersed* condition–a "2-octaves" range–than in the *Unordered-expanded* condition–a "3-or-more-octaves" range; that is, it might reflect an effect of range. However, this interpretation is not supported by other comparisons–e.g., *Original-dispersed* versus *Original-Expanded*–either in the confidence levels or in the proportion of correct tonics. Alternatively, the difference might reflect the influence of some factor other than those examined here, in particular, auditory streaming–see General Discussion.)

Second, the levels of confidence were significantly higher in the *Unordered* condition than in the other conditions (*Original* condition excluded). This implies that pitch proximity, the only factor by which the *Unordered* sequences differed from the remaining distorted sequences, prevented the feeling of tonal ambiguity. Further, given that pitch proximity did not prevent the "loss" of the tonic (in the *Unordered* condition), it also implies that confidence reports (i.e., the feeling of tonal ambiguity/clarity, or un/certainty) and tonic judgments (i.e., key estimation) work differently. That is, that the affective and cognitive mechanisms entailed in these responses, respectively, are at least partially independent. Indeed, the correlation between the (average) levels of confidence each participant reported in each condition and the (average) proportions of CT they identified was $r(210) = .46$, $p < .001$, which means that (after calculating the corresponding coefficient of determination) the 'hits' in the tonic identification task account for only 21% of variation in confidence. Interestingly, the correlation decreased to $r(175) = .22$, $p = .004$, when responses from the *Original* condition were deleted, but remained roughly constant, with a mean $r = .49$ (range $r = .43$ - $.52$; ps < .001), when responses from any of the other conditions were deleted instead. This suggests that when the *Original* sequences were distorted, the convergence of affective and cognitive mechanisms during tonality induction was even weaker.

Finally, Figure 4 (lower panel) shows that the levels of confidence decreased towards the *Unordered-dispersed* and *Unordered-expanded* conditions, but also that only decrements towards the *Unordered-expanded* condition were all significant. This suggested that the effects of pitch class ordering and pitch proximity on participants' confidence were not additive, as they were in the tonic identification task. A two-way repeated-measures ANOVA with Order and Transposition as within-subjects factors confirmed this intuition–thus providing further support to the idea that the affective and cognitive mechanisms involved in tonality induction work differently. This analysis showed significant main effects of Order and Transposition, $F(1, 34) = 75.60$, and $F(2, 33) = 85.96$, respectively; ps < .001, and also a significant interaction between these factors, $F(5, 30) = 38.41$, $p < .001$. To break down this interaction, simple contrasts were performed taking the "original" level of Order and the "none" level of Transposition as controls. These contrasts revealed that the "2-octaves" and "3-or-more-octaves" transposition (compared to "none") lowered participants' confidence significantly more when the "original" instead of the "random" orders (of pitch classes) were given, $F(1, 34) = 25.10$ and $F(1, 34) = 9.51$, respectively; ps < .01. This means that when pitch class ordering was distorted, the effect of pitch proximity on participants' confidence was weakened.

Experiment 2

Experiment 1 showed that pitch class-related information plays an important role in tonality induction, as proposed by the distributional and functional views. Pitch class distribution predicted quite well the tones participants preferred as tonic tone when they were given the *Original* sequences, whereas pitch class ordering predicted quite well how participants' preference varied when they were given the *Unordered* sequences. However, it revealed that pitch-related information plays an important role in tonality induction too, a fact that neither the distributional nor the functional view take into account. Specifically, evidence was found that listeners' preference for a given ("correct") tonic as well as their confidence in the tonic they identify decrease when proximity in pitch between tones is reduced (i.e., when the number of semitones between them increases). This means that pitch proximity contributes to the tonicization of music: first, by facilitating tonality induction; and second, by strengthening the sense of tonality (via the reinforcement of the feeling that a given tone is indeed the tonic).

The results of Experiment 1, however, could have been due to factors other than pitch. In this regard, it must be noted that whenever the pitch class ordering of tone sequences was distorted, the relationships between pitch class-related or pitch-related information and time-related information were distorted too. Specifically, the distortion of ordering affected the way pitches were metrically accented, because of temporal periodicity, and rhythmically accented, because of short-long durational patterns. Consider, for example, the sequences shown in Figure 3: two of the most stable tones in the key of G major, G and B, were metrically or rhythmically accented in (the first beat of) the first and last measure of the *Original* sequence, respectively; however, neither G nor B was accented in the unordered versions of the sequence–in which A was accented instead. Clearly, these temporal distortions could have weakened tonality induction by giving emphasis to pitches that were not tonally structural. Indeed, there is substantial evidence that metric and rhythmic accents serve to highlight a melody's underlying tonal structure (Abe & Okada, 2004; Boltz, 1989, 1999; Boltz & Jones, 1986; Deutsch, 1980; Hershman, 1995; Jones, Johnston, & Puente, 2006; Schmuckler & Boltz, 1994), which raises the possibility that once accents were altered, the sequences' tonal structure was altered too. Moreover, there is evidence that accented tones are better remembered than unaccented tones (Jones, 1976; Jones, Boltz, & Kidd, 1982; Jones et al. 2006; Jones, Moynihan,
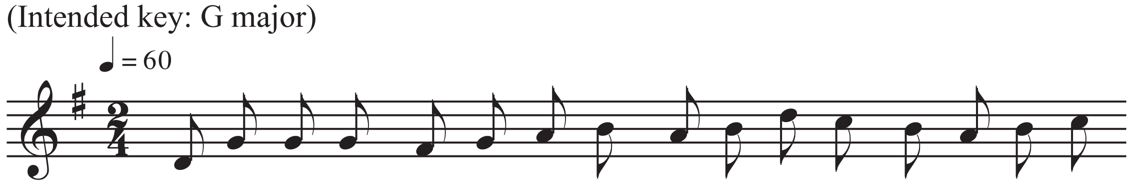
FIGURE 5. G major sequence (*Original* condition) used in Experiment 2: the sequence illustrates the time-related modifications made to the sequences from Experiment 1 (compare Figure 3, *Original* condition).

MacKenzie, & Puente, 2002; see also Boltz, 1999), which raises the possibility that once accents were altered, listeners were simply unable to readily memorize which tone was the tonic.

Thus, although the distortion of pitch-time relationships could not explain why participants' responses varied when pitch class ordering was preserved, it could explain why participants' responses varied when pitch class ordering was distorted: different responses could have been elicited because different tones (i.e., different pitch classes) were metrically and/or rhythmically accented. The potential confounding effect of accents, in turn, severely undermined the idea that pitch class ordering influences tonality induction, proposed by the functional view. Furthermore, it also undermined the idea that during tonality induction the effects of pitch class ordering and pitch proximity are similar, and that pitch proximity prevents the feeling of tonal ambiguity, suggested in Experiment 1. A second experiment was then conducted in which the potential confounding effects of metric and rhythmic accents on tonality induction were eliminated by using isochronous instead of nonisochronous tone sequences as stimuli.

METHOD

Participants. Thirty-five undergraduate students of music from the Faculty of Fine Arts of the UNLP participated in the experiment voluntarily; none of them had participated in Experiment 1. The participants' mean age was 24 years (range: 19-31 years). They had a mean of 6 years of formal training (range: 3-11 years), including ear training, music theory, and music performance, and a mean of 9 years (range: 4-17 years) of informal experience singing or playing an instrument. None of the participants reported having absolute pitch, nor did they report any vocal impediment to sing.

Stimuli and apparatus. The stimuli used in Experiment 2 were the same as in Experiment 1, except that their time-related information was modified in two ways: first, the rhythms of the sequences were replaced by an isochronous rhythm in which there were two tones in each beat; and second, the *tempi* of the sequences were replaced by a *tempo* of 60 beats per minute. Figure 5 shows the original version of the G major sequence used in Experiment 2 to illustrate these modifications. Because of them, in Experiment 2 all the sequences had a steady rhythm in which each tone lasted 0.5 s. The apparatus were the same as in Experiment 1.

Procedure. The procedure was the same as in Experiment 1. (As in Experiment 1, informal post-test interviews suggested that participants were not familiar with the original melodies used as stimuli.)

RESULTS AND DISCUSSION

As in Experiment 1, only responses from test trials were analyzed herein. Again, the first question addressed was which tones the participants preferred as tonic tones when they were given the *Original* sequences. Similarly to Experiment 1, 30 of the 35 participants preferred the intended tonics, which were selected on average 80.0% of the times. Thus, this result showed that even when metric and rhythmic clues were eliminated, the *Original* sequences conveyed the tonics that they were expected to convey. Accordingly, in subsequent analyses the intended tonics for the *Original* sequences were considered as Correct Tonics (CT), whereas other tonics were considered as Incorrect Tonics (IT).

A repeated-measures ANOVA examined the effect of Melodic Condition on participants' preference for CT. The results of this analysis are depicted in Figure 6 (upper panel). As in Experiment 1, the effect Melodic Condition was significant, $F(5, 30) = 29.74$, $p < .001$. Bonferroni post hoc comparisons showed that the proportions of CT were again higher in the *Original* condition than in the other conditions (all *p*s < .001). Next, Friedman's ANOVAs were applied separately to data from each melodic set. Table 2 (upper portion) summarizes the observed results. As can be seen, in five of the six sets (the Eb major set excluded) the average proportion of CT was higher in the *Original* condition than in the distorted conditions, which confirms that participants' ability to find the key tended to decrease

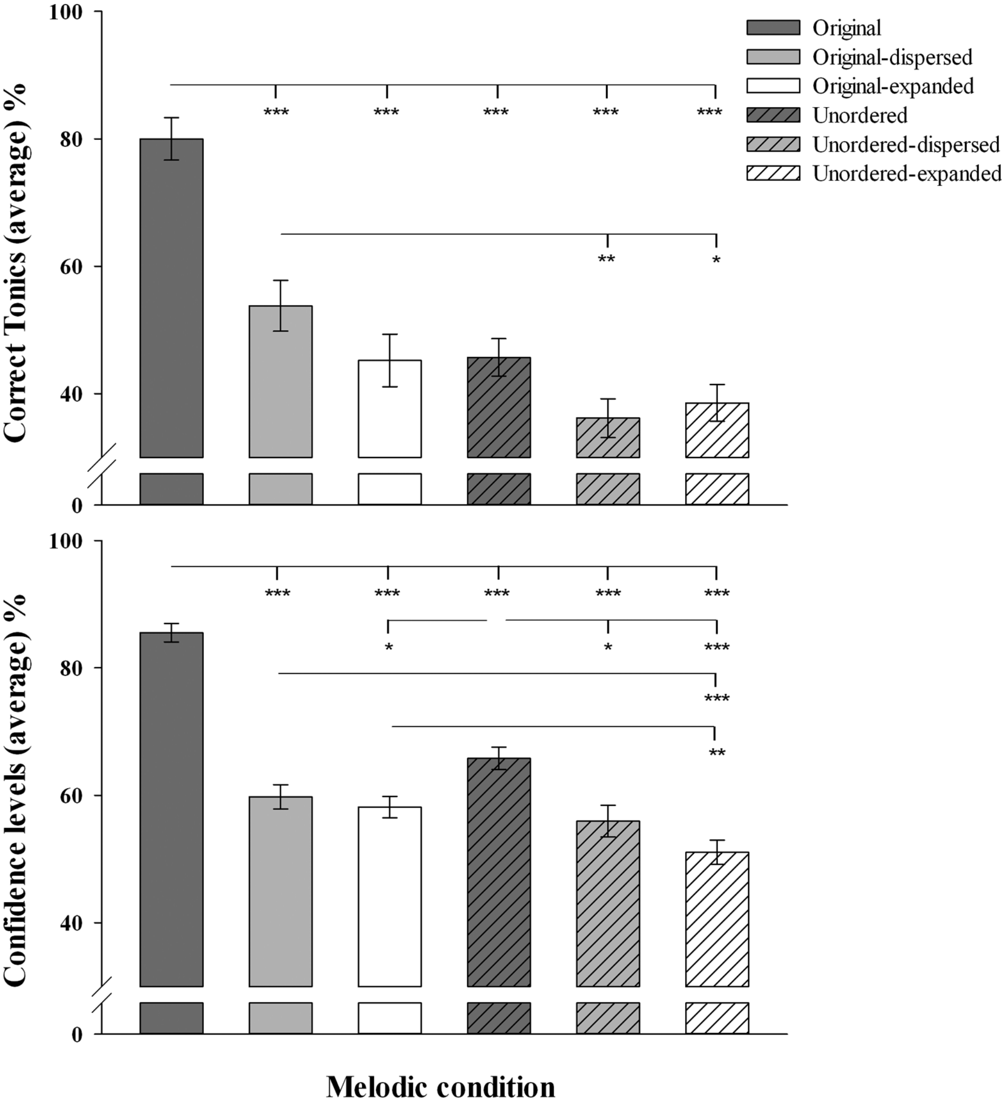

FIGURE 6. Average proportions (in %) of Correct Tonics and Confidence Levels observed in each melodic condition in Experiment 2. Asterisks indicate statistical significance of the difference between conditions (* $p < .05$, ** $p < .01$, *** $p < .001$). Error bars represent SEMs.

when the *Original* sequences were distorted. (Differences between sets were examined by testing the effects of some additional factors–e.g., their tonal strength–as in Experiment 1, though again the results of these analyses were not significant.)

However, Figure 6 (upper panel) shows that, in this second experiment, decrements in participants' ability to identify the CT were unevenly distributed across conditions; beyond the differences between the *Original* and distorted conditions, there were only two other significant differences, one between the *Original-dispersed* and *Unordered-dispersed* conditions, and the other between the *Original-dispersed* and *Unordered-expanded* conditions. This suggested that the effects of pitch class ordering and pitch proximity on tonic selection had not been additive, as in Experiment 1, but interactive. A two-way repeated-measures ANOVA with Order and Transposition as within-subjects factors confirmed the interaction, $F(2, 33) = 15.85, p < .001$. As in Experiment 1 (when confidence ratings were analyzed), this interaction was examined with simple contrasts using the 'original' level of Order and the "none" level of Transposition as controls. The contrasts revealed that the "2-octaves" and "3-or-more-octaves" transpositions (compared to "none") lowered the proportion of CT significantly more when the sequences had their "original"

TABLE 2. Average proportions (in %) of *Correct Tonics* and *Confidence Levels* observed in each Melodic Set as a function of Melodic Condition in Experiment 2.

Correct Tonics

	Melodic Condition					
Melodic Set	Original	Original-dispersed	Original-expanded	Unordered	Unordered-dispersed	Unordered-expanded
G major*	88.57	60.00	71.43	77.14	57.14*	62.86
E major**	65.71	54.29	31.43*	31.43*	11.43***	40.00
Eb major***	77.14	54.29	20.00***	85.71	74.29	68.57
F# minor***	80.00	45.71*	22.86***	14.29***	14.29***	5.71***
D minor***	91.43	57.14**	71.43	11.43***	17.14***	17.14***
C minor**	77.14	51.43	54.29	54.29	42.86	37.14**
Mean	80.00	53.81***	45.24***	45.71***	36.19***	38.57***
SEMs	6.61	8.51	7.68	6.86	7.06	7.23

Confidence Levels

	Melodic Condition					
Melodic Set	Original	Original-dispersed	Original-expanded	Unordered	Unordered-dispersed	Unordered-expanded
G major***	90.61	65.71***	64.90***	75.92**	66.94***	62.86***
E major***	84.08	59.18***	58.37***	65.71***	48.16***	50.20***
Eb major***	85.71	65.31***	59.18***	82.45	69.80***	64.49***
F# minor***	80.82	57.96***	51.84***	47.35***	52.65***	46.53***
D minor***	87.35	57.96***	63.67***	60.41***	60.00***	49.39***
C minor***	83.67	69.39*	64.49***	67.76***	56.33***	53.88***
Mean	85.37	62.58***	60.41***	66.60***	58.98***	54.56***
SEMs	2.42	3.30	3.16	2.97	3.65	3.49

* $p < .05$, ** $p < .01$, *** $p < .001$.

Note: For each Melodic Set, asterisks indicate statistical significance of the effect of Melodic Condition. For each Melodic Condition, asterisks indicate statistical significance of the difference between values in the Original condition and the marked conditions; values in unmarked conditions did not differ significantly from those observed in the corresponding Original condition.

orders, instead of the "random" ones, $F(1, 34) = 9.75$ and $F(1, 34) = 30.86$, respectively; *p*s < .01. Clearly, the simplest explanation for this discrepancy with Experiment 1 is the absence of temporal cues to find the key. However, this explanation is not free from difficulties.

As argued above, the distortion of pitch class ordering in Experiment 1, where temporal structure of stimuli was preserved, could have given salience to pitch classes other than those structural in the intended keys, which finally would have reduced the proportions of CT–in the unordered conditions. This would imply that, consistent with previous studies (e.g., Abe & Okada, 2004; Jones et al., 2002), there were pitch-time interactions across conditions causing the "correct" temporal cues of the original conditions to become "incorrect" cues in the unordered conditions–thus impairing tonality induction. If so, once temporal cues were deleted the proportions of CT in the unordered conditions should increase; further, in the original conditions they should decrease. Interestingly, a comparison between Figures 4 and 6 (upper panels) indicates that the proportions of CT in the unordered conditions were similar (*Unordered* condition) or higher (*Unordered-dispersed* and *Unordered-expanded* conditions) in Experiment 2 than in Experiment 1 (mean difference = -1.9%, 3.8%, and 10.5%, respectively), whereas in the original conditions they were similar (*Original* and *Original-dispersed* conditions) or lower (*Original-expanded* condition) (mean difference = 1.4%, -1.4%, and -9.1%, respectively)–a difference of ±2% is taken arbitrarily as indicating "similarity."[4] Therefore, one may conclude that, in this second experiment, the absence of temporal cues tended to equalize the proportions of CT across conditions, thus eliciting an Order by Transposition interaction.

---

[4] The proportions of CT from both experiments were analyzed using a repeated-measures mixed ANOVA with Melodic Condition as the within-subjects factor and Temporal Cues (two levels: *present* or *absent*) as the between-subjects factor. This analysis demonstrated that the interaction between these factors was small, but significant, $F(5, 64) = 2.52$, $p = .04$. As described in the text, in the original conditions the proportions of CT tended to be higher in Experiment 1 than in Experiment 2, whereas in the unordered conditions the opposite was the case. (However, this did not hold true for the *Original* and *Unordered* conditions–see text.)

However, as the reader may have noticed, within conditions the picture emerging across experiments is problematic: specifically, in the *Original* and *Unordered* conditions the proportions of CT varied in the opposite way to that suggested by the absence-of-temporal-cues explanation. These contradictions, however, are most easily attributable not to a real trend (which would imply, for example, that in the *Original* condition temporal cues inhibited tonality induction), but to random variations in the data. Thus, one may conclude that in the *Original* and *Unordered* conditions there was no difference between experiments. In the *Original* condition, this could have happened because, after all, temporal cues could be "present" in this second experiment too: note that the tonic-triad tones tended to recur periodically in the sequences, every two beats (particularly at the beginning–see Figure 2); further, given that listeners tend to impose "accents" to tones (every two or four indeed) even in perfectly regular sequences (Fraisse, 1982; Potter, Fenwick, Abecasis, & Brochard, 2009; Repp, 2007), periodicity might have triggered this tendency. However, this does not hold for the *Unordered* condition, where pitch class ordering was random. Finally, since the only feature the various *Original* and *Unordered* sequences shared (across experiments) was that their tones were close to each other in pitch, it may be that pitch proximity has compensated both the absence of "correct" temporal cues in the *Original* condition (Experiment 2), and the presence of "incorrect" temporal cues in the *Unordered* condition (Experiment 1). If so, one should conclude not only that the absence of temporal cues elicited an Order by Transposition interaction, but also that pitch proximity regulated participants' need for temporal cues to find the key. Future research should explore these possibilities.

Notwithstanding these differences between experiments, the results of Experiment 2 closely mirrored those of Experiment 1 in three important ways. First, the proportions of CT were again higher in the *Original* condition than in all the other conditions. This confirms that the distortion of pitch class ordering and pitch proximity affects tonality induction. Second, the proportions of CT did not differ significantly from the *Original-dispersed* to the *Original-expanded* condition. Thus, no support was found for the hypothesis that pitch contour influences tonality induction. Finally, the proportions of CT did not differ significantly from the *Original-dispersed* and *Original-expanded* conditions to the *Unordered* condition. This supports the idea that the distortion of pitch proximity and pitch ordering inhibit tonality induction to a similar degree.

Participants' incorrect responses were examined next. As in Experiment 1, the proportions of related and unrelated IT from each participant in each melodic condition were compared using a two-way repeated-measures ANOVA with Tonal Status and Melodic Condition as within-subject factors. (There were three test trials in which three different participants did not respond; i.e., did not sing a tonic. These cases were considered as unrelated incorrect responses.) As expected, the main effect of Melodic Condition was significant, $F(5, 30) = 29.41$, $p < .001$, with fewer errors in the *Original* condition (6.6%) than in the distorted conditions (mean 18.7%; range: 15.4%–21.5%). The main effect of Tonal Status was also significant, $F(1, 34) = 37.77$, $p < .001$, reflecting the fact that related IT (68.0%) were selected more often than unrelated IT (32.0%). However, the interaction between Melodic Condition and Tonal Status was not significant, $F(5, 30) = 1.86$, $p = .13$. Recall that in Experiment 1 this interaction was significant, but not as predicted; differences between related and unrelated errors were larger in the distorted conditions. In light of the present results, it seems that that trend was not psychologically relevant, which would explain why the interaction was not significant in the present data. This leads one to conclude that the sense of tonal relatedness between keys is not substantially altered when pitch class ordering and/or pitch proximity are distorted.

Finally, participants' confidence ratings were analyzed. First, average confidence ratings were entered into a one-way repeated-measures ANOVA with Melodic Condition as the within-subjects factor, as in Experiment 1. Figure 6 (lower panel) depicts the results of this analysis. Again, the effect of Melodic Condition was significant, $F(5, 30) = 63.59$, $p < .001$. Bonferroni post hoc comparisons showed that the levels of confidence were higher in the *Original* condition than in the other conditions (all *p*s < .001). Next, one-way repeated-measures ANOVAs were conducted to examine whether the same pattern of results emerged within each melodic set. The results of these analyses are summarized in Table 2 (lower portion). As may be seen, the distribution of the data within each set largely resembles that found when the average data were examined (particularly in the case of the major-key sets).

Figure 6 (lower panel) also shows that the levels of confidence were again significantly higher in the *Unordered* condition than in the remaining distorted conditions, as in Experiment 1. This finding supports the idea that pitch proximity not only enhances the key-finding process but also prevents the feeling of tonal ambiguity. Further, given that, as in Experiment 1, pitch proximity did not prevent the loss of the tonic (in the *Unordered* condition), it also supports the idea that the affective and cognitive mechanisms involved in tonality induction work differently. In line with this, the correlation between the (average) levels of confidence and the (average) proportion of CT the participants reported in each condition was $r(210) = .50$, $p < .001$, which means that the hits in the tonic identification task account for only 25% of the variation in confidence. The correlation decreased to $r(175) = .24$, $p = .002$ when responses from the *Original* condition were excluded from the analysis, but remained roughly constant, with a mean $r = .53$ (range $r = .50$ - $.57$; $ps < .001$), when responses from any other condition were excluded, which indicates that, as in Experiment 1, the convergence between the affective and cognitive mechanisms involved in tonality induction was weaker when the sequences were distorted.

Finally, to test whether there was an interaction between pitch class ordering and pitch proximity, as there was in Experiment 1, confidence levels were analyzed using a two-way repeated-measures ANOVA with Order and Transposition as within-subjects factors. Both the main effect of Order, $F(1, 34) = 73.18$, $p < .001$, and Transposition, $F(2, 33) = 57.66$, $p < .001$, were significant; the proportions of CT were higher when the sequences had their "original" orders, or "none" transposition. The interaction between Order and Transposition was also significant, $F(2, 33) = 26.82$, $p < .001$. Simple contrasts revealed that the "2-octaves" and "3-or-more-octaves" transpositions (compared to "none") lowered participants' confidence significantly more when pitch classes followed the "original" orders, $F(1, 34) = 42.74$ and $F(1, 34) = 49.62$, respectively; $ps < .001$. This means that, again, the effect of pitch proximity on confidence levels was weaker when pitch class ordering was distorted.

General Discussion

Previous studies on tonality induction have mostly been focused on the effects of pitch class-related information. In this context, two views have become increasingly influential, the distributional view and the functional view. According to the former, tonality induction relies on pitch class distribution (e.g., Bharucha, 1987; Krumhansl, 1990; Temperley & Marvin, 2008; Tillman et al., 2000), whereas according to the latter, it relies on interval class vectors and pitch class ordering (e.g., Brown & Butler, 1981; Brown et al., 1994; Browne, 1981; Matsunaga & Abe, 2005). In contrast, the present study focused on how pitch-related information affects tonality induction. The results indicate that listeners' ability to identify the tonic of a sequence of tones as well as listeners' confidence in the tonic they identify decrease when the distance in pitch between successive tones increases, and vice versa. Moreover, it was also found that the effects of pitch proximity and pitch class ordering on tonality induction are similar and tend to reinforce each other, in the sense that listeners track down the tonic more accurately and confidently when tones are both proximate in pitch and properly ordered in time. Although, strictly speaking, these findings do not refute the hypotheses proposed by the distributional and functional views, they strongly suggest that not only abstract pitch class structures but also concrete pitch patterns must be taken into account in order for tonality induction to be satisfactorily understood.

To the author's knowledge, the present study is the first to demonstrate that pitch proximity has a significant effect on tonality induction. However, a somewhat more striking result was the nonsignificant effect of pitch contour. Indeed, based on conventional wisdom from musical treatises (e.g., Aldwell & Schachter, 2003; Piston, 1941/1959) and some evidence from the psychological literature (e.g., Bharucha, 1984; Cuddy & Lyons, 1981), here both effects were expected to be of similar magnitude. A question that naturally arises, then, is: why pitch contour did not influence tonality induction while pitch proximity did? The simplest answer seems to be: because, in psychological terms, pitch contour largely depends on pitch proximity. This is reflected in two facts. First, whenever pitch contour is distorted (e.g., B4-C5 versus B4-C4), pitch proximity is distorted too, but the opposite is not the case. Therefore, at least in some perceptual tasks it might be enough to pay attention to pitch proximity, disregarding pitch contour, in order for listeners to understand the pitch structure of music. One of those tasks might be tonality induction, which would explain why there was no evidence supporting the influence of pitch contour on either the tonic identification or the confidence level task.

Second, and perhaps more importantly, the very perceptual meaning of "pitch contour" emerges when successive tones are proximate in pitch. Regarding this, Bregman and colleagues (see Bregman, 1990, pp. 417-442; see also Deutsch, 1972) have shown that, unless a schema-governed attentional mechanism is used to reinforce the grouping process, the auditory system prefers to group successive tones that are in the same register (i.e., proximate in pitch) rather than to follow contour trajectories across the registral space. This means, for example, that A3 in the second bar of the *Original-dispersed* sequence shown in Figure 3 may not have been perceived as "going upward to B4," but rather as "belonging to a different group (or auditory stream) from B4" plus "going upward to B3 (in the second beat of the bar)." If so, then listeners would have perceived the sequence not as a single one, but as a set of simultaneously occurring sequences in which the lowest one would be A3-B3-C4-C4. Clearly, this sequence does not seem to convey the key of G major, but rather the key of C major. Interestingly, this grouping-based interpretation would explain not only why pitch contour did not influence tonality induction, but also why pitch proximity did.

Could the results of the present study be due to factors other than those specifically addressed? For example, it is possible that participants' responses may have been affected by their melodic expectations, thus reflecting not the tone they thought to be the most serious candidate as the tonic but the tone they felt to be the most expected. Indeed, in the D minor sequence (which stopped in the leading tone) these two tones could clearly have been the same. Regarding this possibility, however, it is important to bear in mind that, although not experts, participants were musicians; that is to say, they were aware of the notions of tonic, tonality, and alike. Hence, it seems reasonable to assume that they were able to disentangle the problem of "tonalness" from that of "expectedness." In line with this assumption, note that in both experiments the preferred tonics for the G major and Eb major sequences were G and Eb, respectively, and not B and G (in each case, the third scale degree), which, based on previous literature (e.g., Larson, 2004), would have been the most expected tones.

There were two other factors that could have affected participants' responses, range (i.e., the span of pitch covered by a melody) and register (i.e., the absolute location of tones in the dimension of pitch height). Specifically, one may hypothesize that, since tonal melodies are usually constrained to a relatively small range (von Hippel, 2000), it was the unusually large ranges–and not the alteration of pitch proximity, or ordering–that caused listeners to lose the tonic in the dispersed and expanded conditions. Recall, however, that overall participants' responses did not differ significantly in these conditions, despite that ranges were clearly smaller in the former than in the latter. Further, the average proportions of correct tonics in the original melodies (see Tables 1 and 2) were not correlated with the size–in semitones–of their ranges ($ns = 6$, $r = .50$ and $.43$ for Experiments 1 and 2, respectively; $ps > .05$). On the other hand, one may hypothesize that, since extremely low or high tones are more difficult to process (Gelfand, 2010), participants' responses in the dispersed and expanded conditions–where some tones were notably low or high–were due to an effect of register. In fact, listeners' sensitivity to tonal relationships between tones varies across register, being poorer in its extremes (Russo, Cuddy, Galembo, & Thompson, 2007). However, according to the literature the highest and lowest tones used here (C2 and B6, which occurred only once in each experiment) were not as extreme as to be problematic, nor did they fall in the portions of register where tonal relationships were found to be clearly unsafe.

Finally, another question one should address is: how does the present approach to tonality induction relate to the distributional and functional approaches? Regarding this issue, it is important to stress that in the pitch-based approach adopted here the role of pitch class-related information in tonality induction is not denied; rather, it is redefined. Instead of thinking about listeners as computing octave relationships between tones to identify the tonic, they are thought of as identifying a tone as the tonic and then generalizing it (and the remaining tones heard) across octaves. That is, tonal relationships between tones (i.e., tonal structure and prolongations) are thought of as existing primarily within the world of pitches, and only secondarily in the world of pitch classes–once tones are transformed, by octaves, from one register to another (on this argument, see Larson, 1997). Interestingly, there is some evidence supporting this possibility. Specifically, Deutsch (1972) found that listeners are able to use octave generalization to confirm the identity of a tune, but not to recognize it without some prior information, which indicates that pitch class-related information operates at a "second stage" of tune recognition (see also Dowling & Hollombe, 1977: Experiment 2). Given this finding, it seems not unreasonable to hypothesize that the same might occur in the context of other musical functions, such as key identification.

Thus, perhaps the most important aspect of the present pitch-based approach, as opposed to the pitch class-based approaches of the distributional and functional views, is that tonality induction is also conceived as a pattern-matching process–and not only as a template-matching one (cf. Schmuckler & Tomovsky, 2005). Since transitions between pitches are taken into account, tonal patterns are assumed to be encoded in memory (mostly in a relative-pitch code) and later retrieved to assist the key-finding process. Interestingly, both music theory (e.g., Piston, 1941/1959; Schenker, 1935/1979) and psychology (e.g., Larson, 1997–1998; Lerdahl & Jackendoff, 1983; Meyer, 1973; Narmour, 1990; see also Lerdahl, 2001) suggest that in tonal music there is a rather limited number of patterns which, in addition, are varied in a rather limited number of ways. This means that a pattern-matching approach to tonality induction would not imply an overwhelming computational burden, as one might think at first glance. Moreover, there is evidence that even a small part of a musical pattern (e.g., an initial, ascending perfect fourth) may give listeners important clues as to what tone is the tonic (Vos, 1999; Vos & Troost, 1989; see also Meyer, 1973), which further supports the psychological plausibility of the present approach.

In sum, the present study indicates that tonality induction depends not only on pitch class-related information, as suggested by the functional and distributional views, but also on pitch-related information. Specifically, it was found that the sense of tonality is more accurate and robust when successive tones are proximate in pitch. This suggests that tone sequences that typically occur in tonal music are encoded by listeners as musical patterns which, in turn, assist them in finding the key of each new musical piece. Octave equivalence effects would intervene in order to ascertain whether the tonal content of a given sequence can be generalized throughout the musical texture. Finally, it is worth mentioning that, from a broader perspective, tonality induction would be affected not only by information on the sequential arrangement of tones, but also by information on how tones are simultaneously combined, as well as on how they are temporally organized. It seems reasonable to expect that some, if not all, of these kinds of information are to be taken into account in order to understand how tonality induction actually works.

Author Note

This work was supported by grants from the Argentine Research Council (CONICET), and Grant #PICT12-1852 from the National Agency for the Promotion of Science and Technology (ANPCyT) of Argentina.

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 no 680, La Plata (1900), Argentina. E-mail: fernandoanta@fba.unlp.edu.ar

References

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

ALDWELL, E., & SCHACHTER, C. (2003). *Harmony and voice leading*. (3rd ed.) Belmont, CA: Thomson-Schirmer.

ALLEN, D. (1967). Octave discriminability of musical and non-musical subjects. *Psychonomic Science*, *7*, 421-422.

ANTA, J. F. (2013). Exploring the influence of pitch proximity on listener's melodic expectations. *Psychomusicology: Music, Mind, and Brain*, *23*, 151-167.

BHARUCHA, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, *16*, 485-518.

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

BHARUCHA, J. J., & KRUMHANSL, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13*, 63-102.

BHARUCHA, J. J., & STOECKIG, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 403-410.

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

BOLTZ, M. G. (1999). The processing of melodic and temporal information: Independent or unified dimensions? *Journal of New Music Research*, *28*, 67-79.

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

Bregman, A. 1990. *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

BROWN, H. (1987). Tonal hierarchies and perceptual context: An experimental study of music behavior. *Psychomusicology*, *7*, 77-90.

BROWN, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception*, *11*, 371-407.

BROWN, H., & BUTLER, D. (1981). Diatonic trichords as minimal tonal cue-cells. *In Theory Only*, *5 (6 & 7)*, 39-55.

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

BROWNE, R. (1981). Tonal implications of the diatonic set. *In Theory Only*, *5 (6 & 7)*, 3-21.

BURNS, E. M. (1981). Circularity in relative pitch judgments for inharmonic complex tones: The Shepard demonstration revisited, again. *Perception and Psychophysics*, *30*, 467-472.

BURNS, E., M., & HOUTSMA, A. J. M. (1999). The influence of musical training on the perception of sequentially presented mistuned harmonics. *Journal of the Acoustical Society of America*, *106*, 3564-3570.

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.

Costa, M., Fine, P., & Ricci Bitti, P. E. (2004). Interval distribution, mode, and tonal strength of melodies as predictors of perceived emotion. *Music Perception*, *22*, 1-14.

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

Cuddy, L. L., & Lyons, H. I. (1981). Musical pattern recognition: A comparison of listening to and studying tonal structures and tonal ambiguities. *Psychomusicology*, *1*, 15-33.

Deutsch, D. (1969). Music recognition. *Psychological Review*, *76*, 300-307.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception and Psychophysics*, *11*, 411-412.

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

Deutsch, D. (1984). Two issues concerning tonal hierarchies: Comment on Castellano, Bharucha, and Krumhansl. *Journal of Experimental Psychology: General*, *113*, 413-416.

Deutsch, D. (2013). *The Psychology of Music* (3rd ed.). San Diego, CA: Elsevier.

Deutsch, D., & Boulanger, R. C. (1984). Octave equivalence and the immediate recall of pitch sequences. *Music Perception*, *2*, 40-51.

Dowling, W. J., & Hollombe, A. W. (1977). The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception and Psychophysics*, *21*, 60-64.

Elliot, J., Platt, J. R., & Racine, R. J. (1987). Adjustment of successive and simultaneous intervals by musically experienced and inexperienced subjects. *Perception and Psychophysics*, *42*, 594-598.

Fraisse P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). New York: Academic Press.

Francès, R. (1958). *La perception de la musique* [The perception of music]. Paris: J. Vrin.

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 367-400). Oxford, UK: Oxford University Press.

Gelfand, S. A. (2010). *Hearing: An introduction to psychological and physiological acoustics* (5th ed.). London, UK: Informa Healthcare.

Greene, R. L., & Samuel, A. G. (1986). Recency and suffix effects in serial recall of musical stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 517-524.

Halpern, A. R., & Bartlett, J. C. (2010). Memory for melodies. In M. R. Jones, A. N. Popper, & R. R. Fay (Eds.), *Music perception* (pp. 233-258). New York: Springer-Verlag.

Hershman, D. P. (1995). Rhythmic factors in tonality. *Psychomusicology, 14*, 4-19.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, *19*, 1-64.

Huron, D., & Parncutt, R. (1993). An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology*, *12*, 154-171.

Janata, P., & Reisberg, D. (1988) Response-time measures as a means of exploring tonal hierarchies. *Music Perception, 6(2)*, 163-174.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*, 323-355.

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception and Psychophysics*, *32*, 211-218

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology*, *53*, 59-96.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13*, 313-319.

Kallman, H. J. (1982). Octave equivalence as measured by similarity ratings. *Perception and Psychophysics*, *32*, 37-49.

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

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

Krumhansl, C. L. (2004). The cognition of tonality – as we know it today. *Journal of New Music Research, 33*, 253-268.

Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory and Cognition, 11*, 325-334.

Krumhansl, C. L., & Cuddy, L. L. (2010). A theory of tonal hierarchies in music. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception* (pp. 51-87). New York: Springer.

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.

Larson, S. (1997). The problem of prolongation in tonal music: Terminology, perception, and expressive meaning. *Journal of Music Theory*, *41*, 101-136.

Larson, S. (1997-1998). Musical forces and melodic patterns. *Theory and Practice*, *22-23*, 55-71.

Larson, S. (2004). Musical forces and melodic expectations: Comparing computer models and experimental results. *Music Perception, 21*, 457-498.

Lee, J. Y., & Green, D. M. (1994). Detection of a mistuned component in a harmonic complex, *Journal of the Acoustical Society of America*, *96*, 716-725.

Lerdahl, F. (2001). *Tonal pitch space*. New York: Oxford University Press.

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

Marmel, F., Tillmann, B., & Delbé, C. (2010). Priming in melody perception: Tracking down the strength of cognitive expectations. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1016-1028.

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

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

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

Parncutt, R. (2011). The tonic as triad: Key profiles as pitch salience profiles of tonic triads. *Music Perception*, *28*, 333-365.

Pineau, M., & Bigand, E. (1997). Effet des structures globales sur l'amorçage harmonique en musique [Effect of global structures on harmonic priming in music]. *L'Année Psychologique*, *97*, 385-408.

Piston, W. (1941/1959). *Harmony*. London, UK: Victor Gollancz.

Potter, D. D., Fenwick, M., Abecasis, D., & Brochard, R. (2009) Perceiving rhythm where none exists: Event-related potential (ERP) correlates of subjective accenting. *Cortex*, 45, 103-109.

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

Repp, B. (2007). Hearing a melody in different ways: Multistability of metrical interpretation reflected in rate limits of sensorimotor synchronization. *Cognition*, *102*, 434-454.

Russo, F. A., Cuddy, L. L., Galembo, A., & Thompson, W. F. (2007). Sensitivity to tonality across the pitch range. *Perception*, 36, 781-790.

Rosner, B. S., & Meyer L. B. (1986). The perceptual roles of melodic process, contour, and form. *Music Perception*, *4*, 1-39.

Schaffrath, H. (1995). *The Essen Folksong Collection in Kern format* [Computer database] (David Huron, Ed.). Stanford, CA: Center for Computer Assisted Research in the Humanities.

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

Schellenberg, E. G., & Trehub, S. E. (1994). Frequency ratios and the perception of tone patterns. *Psychonomic Bulletin and Review*, *1*, 191-201.

Schenker, H. (1935/1979). *Free composition (Der Freie Satz): Vol. III of New Musical Theories and Fantasies* (Ernst Oster, Trans.). New York: Longman.

Shepard, R. N. (1964) Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, *36*, 2346-2353.

Schmuckler, M. A. (1997). Expectancy effects in memory for melodies. *Canadian Journal of Experimental Psychology*, *51*, 292-305.

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

Schmuckler, M. A. (2010). Melodic contour similarity using folk melodies. *Music Perception*, *28*, 169-193.

Schmuckler, M. A., & Boltz, M. G. (1994). Rhythmic and harmonic influences on musical expectancy. *Perception and Psychophysics*, *56*, 313-325.

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

Sergeant, D. (1983). The octave–percept or concept. *Psychology of Music, 11*, 3-18.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19*, 110-120.

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

Surprenant, A. M. (2001). Distinctiveness and serial position effects in tonal sequences. *Perception and Psychophysics*, *63*, 737-745.

Temperley, D. (2008). A probabilistic model of melody perception. *Cognitive Science*, *32*, 418-444.

Temperley, D., & Marvin, E. W. (2008). Pitch class distribution and the identification of key. *Music Perception*, *25*, 193-212.

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

Tillmann, B., Janata, P., Birk, J., & Bharucha, J. J. (2003). The costs and benefits of tonal centers for chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 470-482.

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

von Hippel, P. (2000). Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals. *Music Perception*, *17*, 315-327.

Vos, P. G. (1999). Key implications of ascending fourth and descending fifth openings. *Psychology of Music*, *27*, 4-17.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception*, *6*, 383-396.