Memory of a Tonal Center After Modulation

Morwaread M. Farbood
New York University

This study examines how long the percept of a tonal center is retained in memory following a modulation to a new key, and how harmonic context in the new key area affects recall of the original key. In Experiment 1, musically trained listeners ($N = 50$) were asked to rate perceived harmonic tension while listening to chord sequences that consisted of three parts: the first section established an initial key, the second section modulated to a new key, and the last section modulated back to the original key. The duration of the new key section ranged from 3 to 21 seconds. The tension slopes following the modulations indicated a gradual decay in the memory of the previous key as the length of the new key section increased. When sequences lacked cadences, traces of the initial key appeared to persist longer. In Experiment 2, musically trained listeners ($N = 31$) were asked to rate harmonic tension while listening to sequences with longer timescales of up to 45 s in a new key area. Additionally, responses to "closed" modulations, which returned to the original key, and "open" modulations, which departed from both the original and new keys, were compared. The combined results of Experiments 1 and 2 showed that the original key was retained in memory after 15-20 s in a new key. However, there was not enough evidence to conclude it persisted beyond 20 s.

Received: December 30, 2012, accepted November 10, 2015.

Key words: modulation, memory, tonal induction, tonal closure, cadences

The concept of tonal closure has long been regarded by music theorists as an aspect of form that has both structural and aesthetic significance. Tonal closure can occur at multiple levels of musical form and even across multiple movements within a work. However, the term is most frequently used in reference to a piece beginning in a designated key and ending—usually articulated by a cadence—in the same key. Typically, there are also intervening excursions to other key areas before a return to the original key. Empirical studies that have examined the perception of large-scale tonal closure have, for the most part, led to the conclusion that the memory of an opening key is limited in duration if there is a modulation to another key. This has cast doubt on the psychological reality of listener recognition of tonal closure since cognitive constraints clearly limit the ability of listeners to retain a previous key in memory. Although some work has been done on how home keys can influence tonal perception after modulation (Cuddy & Thompson, 1992; Thompson & Cuddy, 1992), no prior work has examined the precise time course for how long a previous key is retained in memory after a modulation and what factors affect the duration of retention. This study explores specific timescales and examines the effect of harmonic context on recall of a previous key.

Cook (1987) was the first to examine large-scale tonal closure from a psychological perspective. Subjects in his study listened to six solo piano excerpts from the classical repertoire in their original form as well as in a modified form that did not end in the original key. Cook asked participants to compare the two versions and indicate which they preferred in terms of expressiveness, coherence, pleasure, and sense of completion. The results indicated that listener preferences for the original were only clear in the case of the shortest excerpt. This short excerpt was under a minute in length, although the exact duration was not specified—it appeared to be around 30 seconds long with a sudden modulation occurring precisely midway through the excerpt. These findings cast serious doubt on the assumption by theorists that listeners recognize large-scale tonal closure and that it serves an important aesthetic function.

Karno and Konečni (1992) explored how structural changes to a Mozart symphony movement affected aesthetic evaluation. They scrambled the movement by section, creating versions that did not begin and end in the same key. Despite these significant alterations, subjects did not prefer the original version over the scrambled ones. Karno and Konečni's study, like later ones that ask subjects to provide aesthetic judgments of scrambled music or piece together musical segments to form a coherent whole, generally concluded that large-scale tonal closure has little to no effect on aesthetic preference (Granot & Jacoby, 2011; Lalitte & Bigand,

Music Perception, volume 34, issue 1, pp. 71–93, issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2016.34.1.71

2006; Tillmann & Bigand, 1996; Tillmann, Bigand, & Madurell, 1998).

Marvin and Brinkman (1999) noted that although Karno and Konečni's study, and others in a similar vein, might have used stimuli that disrupted tonal coherence, they mainly focused on perceived expressiveness. Although large-scale tonal closure did not appear to influence aesthetic perception in these studies, that did not necessarily imply that listeners were unable to *perceive* large-scale tonal closure. Marvin and Brinkman addressed the question of tonal closure perception more directly by asking subjects to explicitly state whether they thought musical excerpts started and ended in the same key. In their first experiment, musicians were asked whether they thought the starting and ending keys were the same for several piano and orchestral excerpts. Results showed that subjects answered this question correctly at levels above chance, particularly in the case of musicologists and theorists. In a second experiment, musicians heard MIDI performances of six Handel keyboard compositions in original and modified forms. The modified forms had rearranged sections that resulted in different ending keys. In this case subjects were unable to discern whether the final key was the same as the beginning key at levels above chance.

Although these results seemed to confirm Cook's (1987) observations, Marvin and Brinkman reflected that perhaps one problem with their experimental design was that participants might not have been able to discern the initial key and were therefore unable to make an informed judgment. They also surmised that inaccuracy might have been due to listeners using other structural cues to make judgments when such cues were not reliable indicators of tonal closure. Questions pertaining to the time course of tonal closure also remained unanswered since the actual time elapsed between the modulations was not systematically examined. Even two pieces that are of the same length can have modulations that occur at very different points in time, resulting in varying levels of recall difficulty.

Woolhouse, Cross, and Horton (2016) explored structural similarities in music and language by asking listeners to rate the level of closure following modulations to new key areas of varying lengths. Although they did not specifically examine timescales of retention, they did note that listeners were able to hold the original key in memory for "over 10 seconds." Woolhouse et al.'s (2016) and Cook's (1987) results seem to point toward a minimum recall time of 10 to 15 seconds (the latter value being a conservative estimate from the Cook study).

There are other studies employing indirect methods that have provided additional evidence for the perceptual limitations of tonal closure. These include empirical explorations of Lerdahl's (1996, 2001) tonal tension model. Lerdahl's model easily lends itself to empirical evaluation because it provides quantitative measures of tonal tension. The model consists of four distinct components: (1) a representation of hierarchical event structure, (2) a model of tonal pitch space and the distances between chords within the space, (3) a treatment of dissonance (primarily psychoacoustic), and (4) a model of melodic/voice-leading attractions. The hierarchical component of the formula is based on the prolongational reduction described in Lerdahl and Jackendoff's (1983) generative theory of tonal music (GTTM). The hierarchies described in GTTM represent a final-state mental representation of a piece of music as processed by an idealized listener. It is not intended to be a real-time model, and as such, it does not have a component representing memory decay or interference.

Lerdahl's model has been tested extensively and shown to be highly successful in predicting listener judgments of harmonic tension (Bigand, Parncutt, & Lerdahl, 1996; Lerdahl & Krumhansl, 2007). Others have questioned the unqualified application of its hierarchical aspects to real-time listener perception; in particular, Bigand and Parncutt (1999) concluded that tension was only weakly influenced by global harmonic structure and was determined more directly by local cadences. Although the influence of tonal hierarchies was essential in describing shorter excerpts in their study, the results suggested that a strict application of global hierarchical structure for calculating tonal tension values does not accurately predict listeners' responses to sustained key changes. Lerdahl's model predicted that listeners would hear an entire section in a new key at an elevated tension level from the previous key while the experimental data indicated that the tension level dropped fairly quickly after the new key was established. Bigand and Parncutt (1999) theorized that there is a "reset" of sorts that occurs when a new key is established. They argued that musical events are perceived through a short perceptual sliding window where events perceived at a given time are negligibly influenced by events outside the window. Despite their conclusions that hierarchical structures are not significantly influential at a global level, they still acknowledged that Lerdahl's model was the most effective of the several models they were testing. Furthermore, a later study by Lerdahl and Krumhansl (2007) using some of the same stimuli showed that Lerdahl's model fit the tension responses more closely. These differences might have been due to a methodological issue since Bigand and Parncutt defined tension as a "feeling that there must be a continuation of the sequence," while Lerdahl and Krumhansl did not provide their subjects with a specific definition of tension.

Regardless of issues surrounding the global influence of tonal hierarchies (in the sense of prolongational structures), there seems to be little doubt that they contribute to harmonic tension perception in shorter time spans. In general, tonal hierarchies are essential to music processing for both musicians and nonmusicans across relatively short temporal windows of 10-15 s (Tillmann & Bigand, 2004). This has been repeatedly demonstrated in numerous harmonic priming experiments as well (Bharucha, 1987; Bharucha & Stoeckig, 1986, 1987; Bigand & Pineau, 1997; Bigand, Tillmann, Poulin-Charronnat, & Manderlier, 2005; Tillmann & Bigand, 2001; Tillmann, Bigand, & Pineau, 1998; Tillmann, Janata, Birk, & Bharucha, 2008).

Farbood (2010) examined how memory limitations might affect perception of hierarchical tonal structures. In this context, the addition of a memory decay component to Lerdahl's (2001) tonal tension model was proposed. The decay component decreased the additive tension factor contributed by the higher-level branches in the prolongational hierarchy. For example, according to Lerdahl's model, if a modulation occurred to a distant key, all of the events in the new key area would inherit the harmonic distance value between the two keys; this would elevate the tension values of the entire new key section regardless of time elapsed following the modulation. Farbood analyzed continuous tension responses using regression analysis that took into account various parameters including harmonic tension, melodic contour, and onset frequency. Descriptions of how these features changed over various time spans ranging from 0.25 to 20 s were used in an attempt to identify the best predictors of the general tension curve. The results indicated that change in harmony best fit the tension data when the time differential was between 10-12 s, while other features best fit the data at a time differential of around 3 s. This suggested that the memory of tonal regions is retained for a considerably longer period of time than is the case for other musical structures such as rhythm and melodic contour. These results are in accord with more general perspectives on memory such as Craik and Lockhart (1972), who theorized that the higher the level of information abstraction, the longer its persistence in memory. Deutsch and Feroe (1981) surmised that this general theoretical framework might apply to music as well. Although there are hierarchical aspects of both rhythmic/metrical structure and pitch contour, the process of tonal induction arguably requires higher-level cognitive abstractions than either (Janata et al., 2002; Koelsch, Gunter, Schröger, & Friederici, 2003).

The consensus that arises from prior work is that there are memory constraints that limit listeners from perceiving large-scale tonal closure. However, results from previous studies do not provide a complete picture of the time course of this process. The goal of this study is to better understand the durational boundaries for how long a tonal center is retained in memory after a modulation to a new key and how context might affect recall. The methodological approach taken here is to examine continuous tension responses to modulating harmonic sequences in which the time spans and local harmonic progressions of the new key area are systematically varied.

General Method

Tension judgments are an effective way of evaluating tonal perception in real time; previous theoretical and empirical investigations have shown that tonal modulations correspond to sharp increases in tension (Bigand & Parncutt, 1999; Farbood, 2012; Lerdahl, 2001; Lerdahl & Krumhansl, 2007). As such, the general idea behind the stimulus design was to feature chord progressions containing two modulations: the first from an established initial key to a new key, and then from the new key back to the original key. The period of time spanned by the new key and the type of local progression in the new key would be manipulated as independent variables.

There were two methodological concerns that needed to be addressed before the experimental design was finalized. (1) With regard to stimulus construction, it is not possible to completely divorce harmonic and melodic changes, since any change in harmony would necessarily result in a change in one or more pitch contours of the outer or inner voices of a chord. Since melodic contour changes have a significant effect on tension perception (Bigand et al., 1996; Farbood, 2012; Granot & Eitan, 2011; Krumhansl, 1996; Nielsen, 1987), it was imperative to control for this. (2) With regard to the task description, it was unclear whether listeners should be asked to rate "harmonic stability" or "harmonic tension." The idea was to find the best task that elicited the clearest responses to *tension generated by harmony alone* as opposed to tension generated by pitch height changes or other possible factors such as anticipation due to repetition.

Informal pilot studies were conducted to resolve these two issues. They involved 3-5 subjects, all of whom were musicians. Several strategies were explored to mask the effect of voice leading in the chord progressions:

FIGURE 1. Some examples of the different versions of stimuli tested in pilot studies.

(1) varying timbre by using sustained (strings) versus less sustained (piano) instrument sounds; (2) adding rests between chords; (3) altering chord voicing—this ranged from standard four-part harmonizations to unorthodox voicings that filled in chords at different registers; (4) changing the loudness of certain tones to undermine voice leading; (5) arpeggiating chords. A few versions of chord progressions that were tested in the pilot studies are shown in Figure 1. None of these strategies were able to solve the problem—melodic contours of both inner and outer voices could not be effectively masked. Sustained timbres worsened the effect and rests appeared to disrupt the continuity of the progressions without lessening the voice-leading effects. Close voicings in different registers, loudness alterations of certain pitches within chords, and arpeggiations were likewise ineffective in neutralizing the problem. This issue seemed to go beyond simple acoustic effects. Experienced listeners of Western tonal music appear to project voice-leading patterns onto typical chord transitions even if they are not overtly present. This is presumably due to the schematic voice-leading patterns learned implicitly and further reinforced through music theory knowledge. In conclusion, the only viable solution was to balance the melodic effects by employing two versions of all chord progressions, each with opposing melodic contours.

Given the results of the pilot testing, the decision was made to render the final version of the stimuli in piano timbre and arpeggiate the chords in a manner idiomatic to piano technique. In addition to sounding more ecologically valid, the arpeggiation was helpful in alleviating the effect of awkward voice leading that was sometimes unavoidable in the process of composing stimuli that strictly adhered to the opposite-contour design. Furthermore, piano sounds were the most generic and least distracting MIDI timbre that could be utilized; when subjects are familiar with the timbre of the stimuli, it does not disturb their concentration on the task (Auhagen & Vos, 2000).

The two tasks of rating harmonic tension versus harmonic stability were explored in the pilot studies as well. Some subjects participated in only one of the rating tasks, and others participated in both. The latter group reported having no problems understanding the distinction between the two tasks. Examination of these preliminary data indicated different types of responses for the two tasks. In the case of stability, modulation would elicit indication of loss of stability, but any cadence following a modulation would tend to result in a maximum stability rating at the point of resolution regardless of how closely the cadence followed an abrupt modulation to a remote key. In other words, the stability task resulted in judgments that were deemed highly local, making it impossible to evaluate how long the previous key was retained in memory. The tension ratings, on the other hand did not reflect this effect: although cadences would result in decreases in tension, they appeared to be tempered by distance from the previous key area as opposed to resulting in a "zero" rating for any authentic cadence. Given these results, it was determined that participants in the actual experiment would rate tension, not stability.

Two additional issues addressed in the pilot experiments concerned the key distance of the modulations and the tempo of the harmonic changes. In the case of harmonic tempo/rhythm, several different tempos ranging from chord changes every 4 seconds at the slow end to chord changes every 500 ms were explored. Chord changes every 1.5 s were deemed a good compromise between tempos that were so ploddingly slow that continuity was lost between chords to transitions that were too quick to allow subjects to respond to individual chord changes.

Both distant and close modulations were explored as well as the length of the modulatory passage. Briefer modulations appeared to be more effective in isolating a "critical" moment for later data analysis. Thus the decision was made to use a dominant seventh chord in the new key as the single transition chord. As for the distance of the modulation, previous work has indicated that modulations to more distant keys take longer to cognitively process than closer keys (Krumhansl & Kessler, 1982). A compromise was made to modulate to a key that was somewhere in the middle between the closest possible modulation (one step on the descending circle of fifths) and the farthest possible modulation (seven steps). The final version of the stimuli featured modulations that were three steps away on the circle of fifths from the original key.

Experiment 1

METHOD

Participants. Although nonmusicians might be able to successfully perform a harmonic tension rating task, the term "harmonic tension" was considered too technical a term for participants completely untrained in music. Therefore, only musically trained subjects were recruited for the study. However, the participants did span a fairly wide range of musical experience. There were a total of 50 participants, most of whom were undergraduate and graduate students at New York University, mean age 23.66 years ($SD = 6.39$), 27 male, 23 female. Subjects had an average of 9.66 years of formal training on a primary musical instrument ($SD = 5.50$) and an average self-rank in instrumental skill level of 3.69 ($SD = 0.96$) on a scale of 1-5. Average number of years of college-level music theory training was 1.68 ($SD = 1.64$) and the mean overall self-ranked music training level was 3.38 ($SD = 1.13$), where 0 = *no training* and 5 = *professional-level training*. Five subjects reported having absolute pitch. Data from these subjects were included in the analysis because previous work on large-scale tonal closure perception has shown that absolute pitch possessors do not recall the original key more accurately than non-absolute pitch possessors (Marvin & Brinkman, 1999).[1]

Stimulus materials. The stimuli were arpeggiated chords progressions, rendered in MIDI piano timbre, consisting of three sections: Section 1 in the original key, Section 2 in the new key, and Section 3 in the old key. The durations of Section 2 were 0, 3, 7.5, 13.5, and 21 s. The shorter (3-13.5 s) versions of Section 2 were all subsequences of the longest 21 s version. Three types of harmonic contexts were explored by varying the chord progressions in Section 2. Type I sequences consisted of clearly functional, conventionally tonal progressions in the new key, with clear cadences. Type II sequences consisted of meandering and unpredictable tonal progressions without any clear cadence in the new key. Type III sequences consisted of a single repeated tonic chord a tritone distant from the original tonic. Sections 1 and 3 were identical within each type and identical across Types I and II. Two versions of each chord progression with opposite melodic contours were composed to control for the perceptual influence of pitch height. All of the chord progressions composed for the experiment are shown in Figure 2.

Procedure. Participants were seated in front of a computer and were presented the stimuli over Sennheiser HD 650 headphones in a sound-isolated chamber. They were asked to indicate changes in harmonic tension while listening to the stimuli by moving a horizontal slider on a MATLAB graphical user interface that used Psychophysics Toolbox Version 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997) for audio playback; slider values recorded ranged from 0-100. Subjects were given three practice trials featuring chord progressions that

[1] In an analysis not included here, mean responses from the five AP subjects were compared to the 45 non-AP subjects and were found to be very similar.

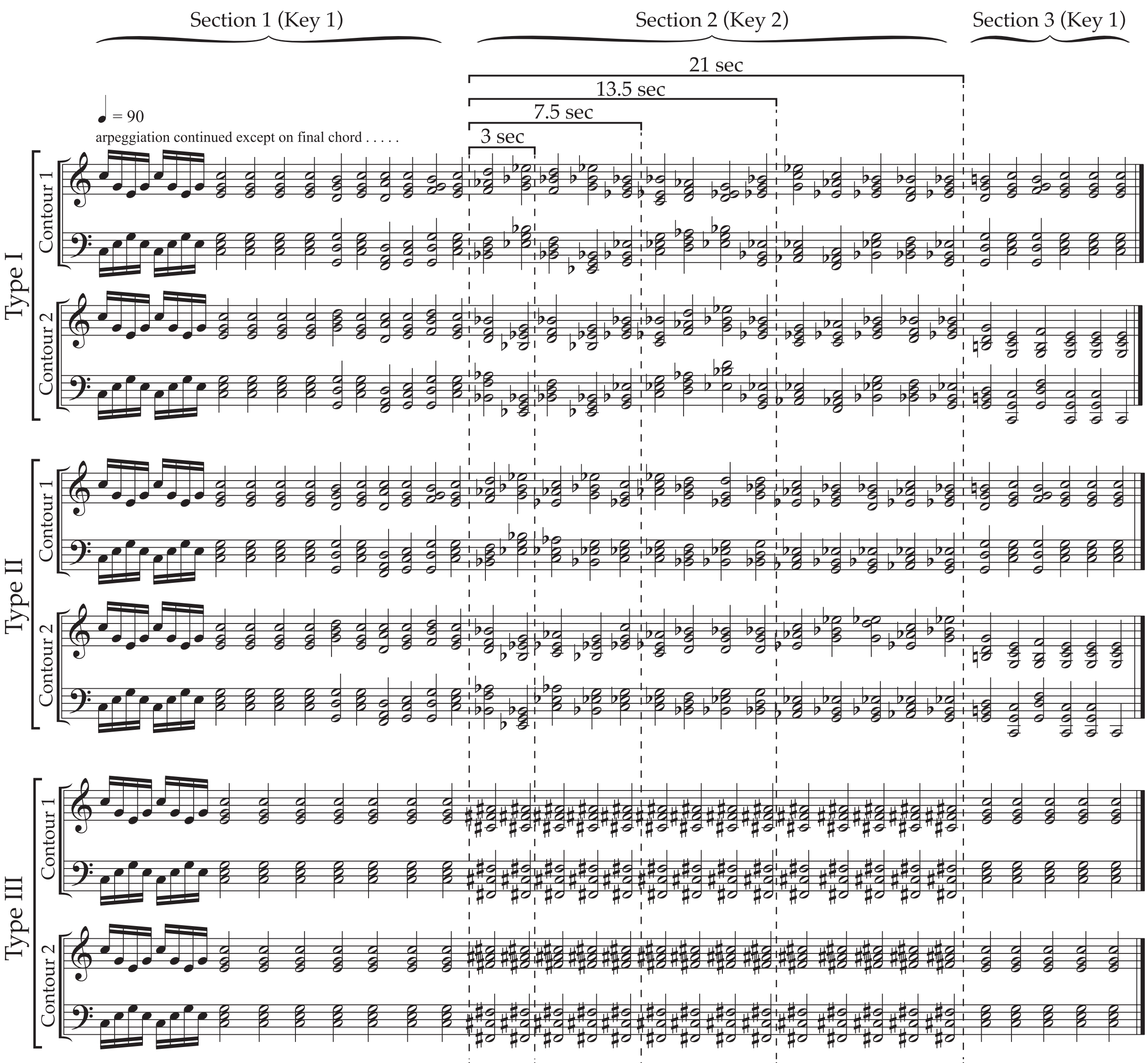

FIGURE 2. The harmonic sequences composed for Experiment 1.

were different from the main stimuli. The practice trials were followed by the recorded trials, in which each of the 30 sequences (3 types × 2 contours × 5 durations) were presented twice in pseudorandom order. Randomization was constrained so that no sequence was followed by another sequence of the same type. All sequences were also transposed pseudorandomly to any of 23 possible transpositions: the original version (as shown in Figure 2), 1 to 11 semitones up from the original, or 1 to 11 semitones down from the original. All transpositions had to be at least three steps on the circle of fifths away (descending or ascending) from the key of the previous sequence and alternated between the ranges spanning the octave below and the octave above the untransposed version.

Data processing. The data analysis focused on two particular moments in each trial: the change in tension immediately following the arrival in the new key and the change in tension immediately following the return to the original key. In other words, the data were analyzed by looking only at the sections of the continuous responses immediately following modulations. The slopes of those intervals, as determined by a linear fit of the region, were calculated for each response.[2] The analyzed sections consisted of 3 s of continuous data following each modulation for Type I and II sequences and 1.5 s for Type III sequences. In the latter case, there was no transition chord to the new key, so the length of the examined section was the duration of a single chord. For Type I and II sequences, it was important to include the response to the new tonic, not just the dominant seventh transition chord, thus the examined section spanned the duration of two chords. This provided enough lag time from the onset of the new tonic to ascertain an immediate reaction.

The idea behind this analysis approach was that the tension slopes following the second modulation (back to the original key) were indicative of how well the first key was retained in memory: a negative slope indicated that the prior key was recalled, since a decrease in tension meant the original key was still the primary key context; a positive slope indicated that the original key had been forgotten, with the new key having replaced any trace of the original key in working memory. All data were normalized within subject to zero mean and unit standard deviation (*z*-score) to account for differences in subjects' rating ranges and response styles.

RESULTS

Mean normalized tension responses for each type and duration are shown in Figure 3. The dashed vertical lines indicate section ends and modulation beginnings (except in the case of the topmost graph, where Section 2 is 0 s and thus the corresponding stimuli have no modulations). The gray highlighted areas indicate the regions where slope values were extracted, and the dotted lines indicate standard error above and below the mean tension responses.

The first step was to examine the average slopes following the initial key change from Section 1 to Section 2 for all of the modulating sequences. As expected, these values were positive for all types: Type I, $M = 0.03$, $SD = 0.02$; Type II, $M = 0.03$, $SD = 0.02$; Type III, $M = 0.07$, $SD = 0.05$. The close match in means and standard deviations, particularly for Types I and II, was not surprising since there were no differences between sequences up through the first key change. These slope values constituted a baseline for perceiving a modulation from an established key to a new key.

The next step was to examine the responses to the second modulation back to the original key for all Section 2 durations, including cases where there was no Section 2 (0 s duration, thus no modulation). The mean slopes at the return to the original key grouped by duration and combined across types are shown in Figure 4. A two-way, repeated-measures ANOVA was performed on the slopes corresponding to the start of Section 3.[3] The main effects of type, $F(1.21, 59.34) = 5.08$, $p < .001$, $\eta_p^2 = .09$, $\eta_G^2 = .03$, and duration, $F(2.99, 146.61) = 13.85$, $p < .001$, $\eta_p^2 = .22$, $\eta_G^2 = .05$, were significant (Greenhouse-Geisser corrected). There was also a strong interaction effect, $F(4.30, 210.79) = 5.98$, $MSE = 0.41$ $p < .001$, $\eta_p^2 = .11$, $\eta_G^2 = .04$, necessitating a look at response profiles individually by sequence type.

The slope values increased as the duration of Section 2 increased; for the 3, 7.5, and 13.5 s conditions, there were negative mean slopes at the return to the original key, while for 21 s, there was a positive mean slope. These results suggested that by 21 s, memory of the original key was substantially weakened. However, the mean slope for the longest duration was only marginally positive and did not come close to the magnitude of the positive slopes corresponding to the initial modulation occurring at the boundary between Sections 1 and 2. This suggested that although the memory of the initial key had faded either due to decay or interference, there were still traces remaining even after 21 s in the new key. A closer look at responses by type would yield a more nuanced story.

Harmonic context. Due to the significant interaction factor, the simple effects of duration for each type were examined individually, all of which were significant (see Table 1 for ANOVA results). The mean slope values broken down by Section 2 duration for each type are shown in Figure 5, and the results of Tukey-Kramer post hoc multiple comparisons are shown in Table 2. Type I (functional harmony) and Type II (meandering harmony) results for the 3 s case were similar, showing negative values in all cases. The data clearly indicated that a short foray into a new key area resulted in a clear decrease in tension when returning to a prior key. On

[2] In addition to slope, the magnitude of tension change over each interval was calculated for Experiments 1 and 2. The subsequent statistical analyses, described in the main text for slope, were also applied to the magnitude change values. These analyses are not included here because the results are very similar to those reported for slope.

[3] Before performing the repeated-measure ANOVA, trials with identical stimuli were averaged. In the case where Section 2 was 0 s, the Type I and II sequences were identical. This meant there were actually four data points per subject in the 0 s case. In all other cases, there were two trials per sequence. In addition to the ANOVA, a linear mixed-effects model was used to analyze the raw data: type and duration were included as fixed effects (duration as a continuous variable) and subject and stimulus as random effects. This analysis is not included here because the results were very similar to the reported ANOVA. In all subsequent analyses for both Experiments 1 and 2, linear mixed-effects models were used in conjunction with ANOVAs; likewise those results are not reported to avoid redundancy.

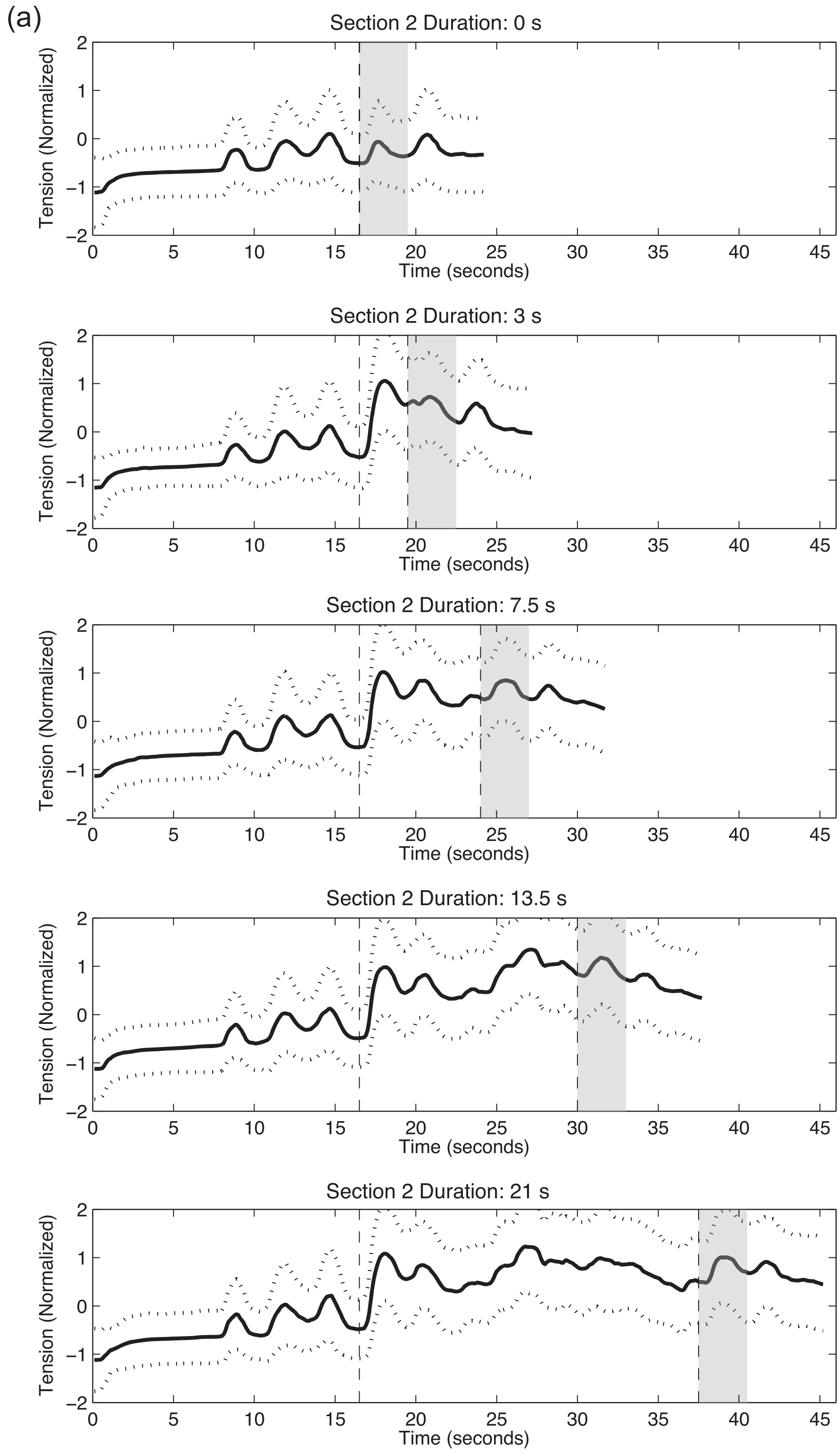

FIGURE 3. Experiment 1 mean tension responses. The dotted lines above and below the graphs indicate estimated standard error below and above the mean; the vertical dotted lines mark the start and end points of Section 2; the shaded regions cover the areas used to extract slope values. (a) Type I sequences (b) Type II sequences (c) Type III sequences.

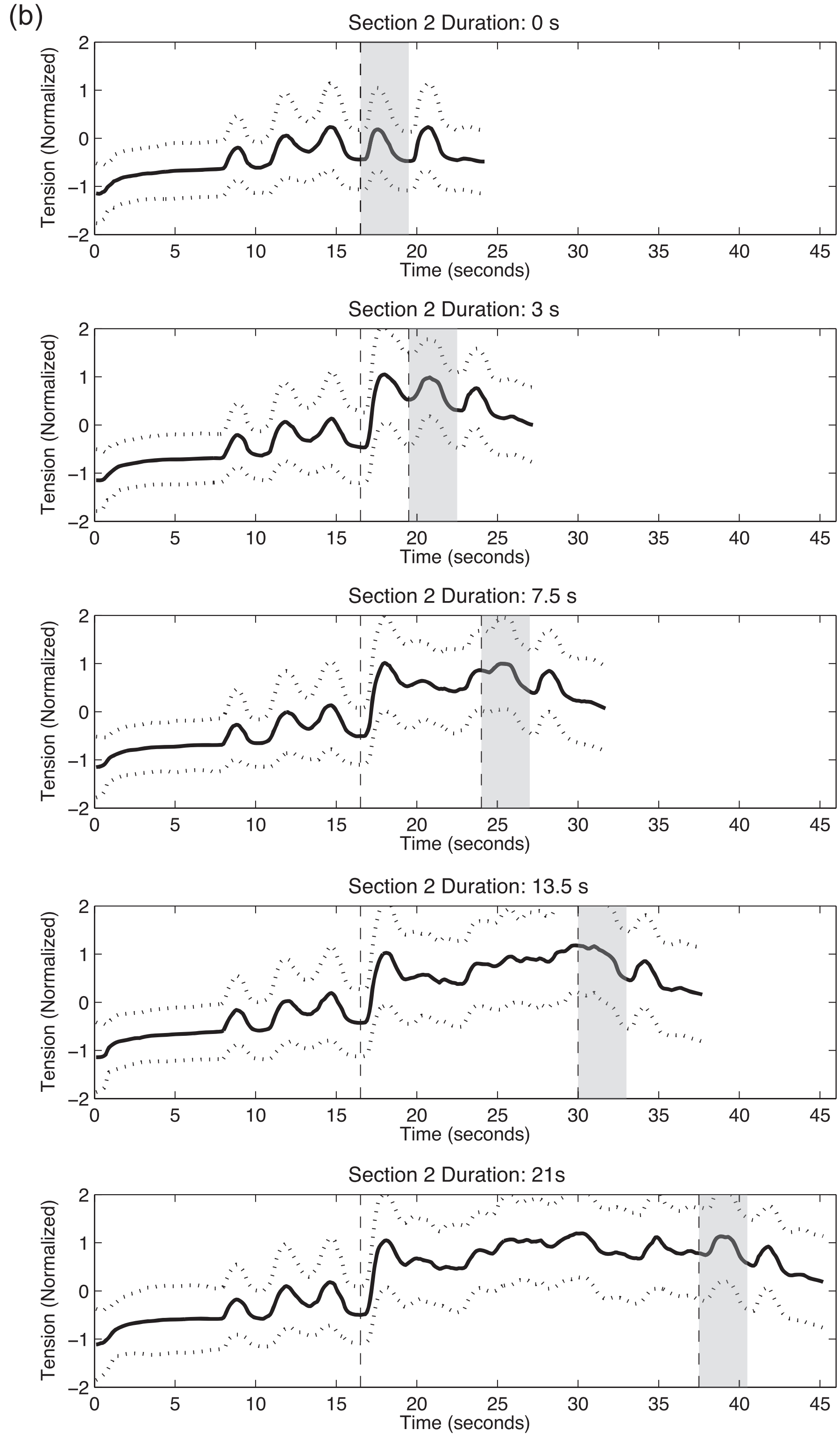

FIGURE 3. [Continued]

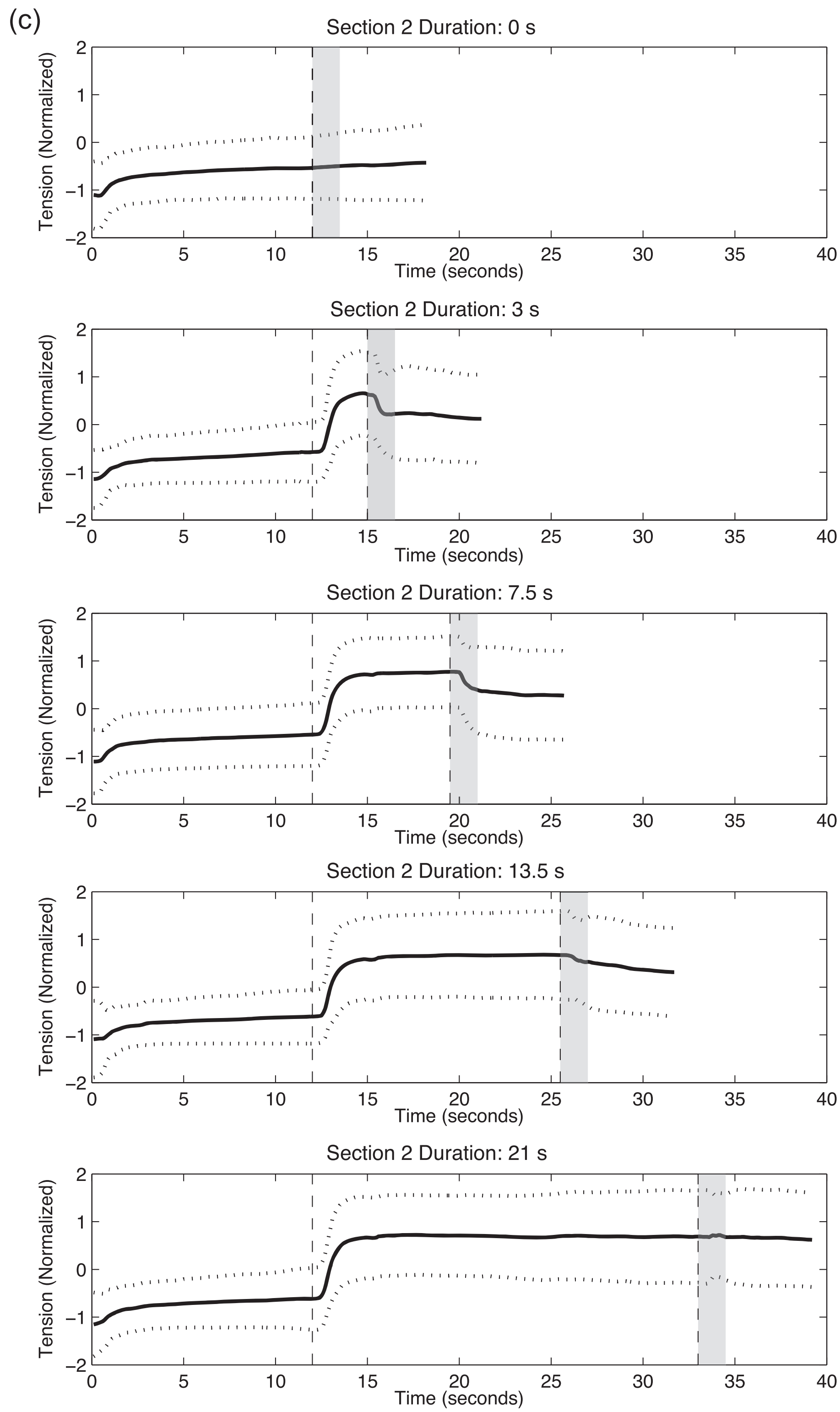

FIGURE 3. [Continued]

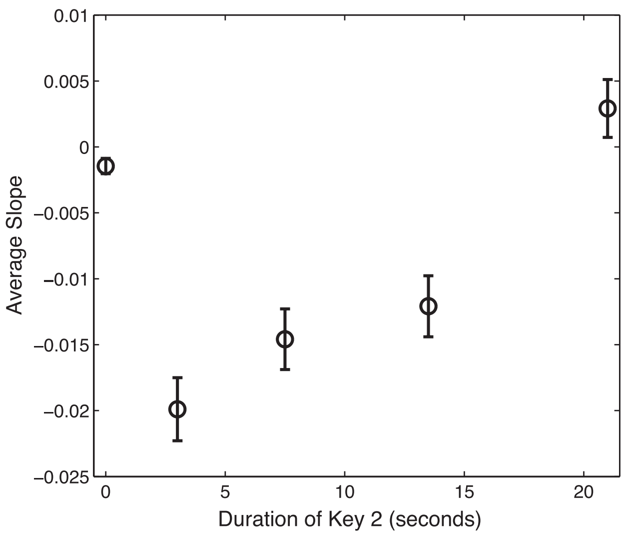

FIGURE 4. Experiment 1 results. Mean tension slopes in response to the modulation back to the original key across all types. Error bars indicate estimated standard error.

TABLE 1. *ANOVA Results for Experiment 1: Simple Main Effects of Section 2 Duration.*

Sequence Type	*F*	*df1*	*df2*	η_p^2	η_G^2	*p*
Type I	11.86	3.54	173.30	.19	.13	< .001
Type II	8.53	3.51	171.88	.15	.10	< .001
Type III	8.41	3.00	146.37	.15	.07	< .001

Note. All results are Greenhouse-Geisser corrected.

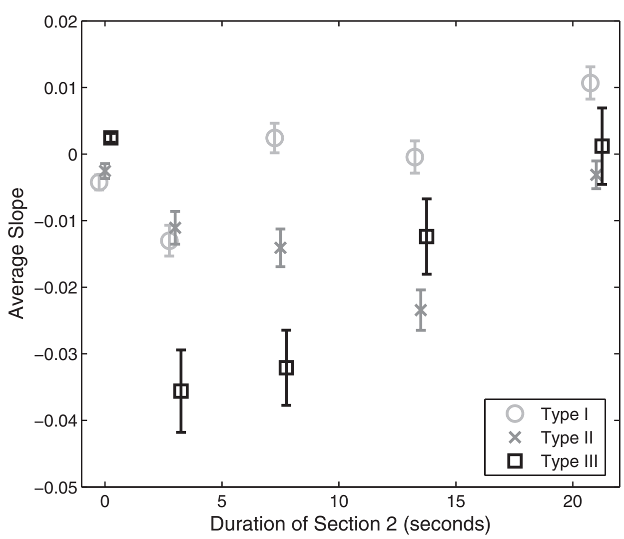

FIGURE 5. Experiment 1 results by type. Mean tension slopes in response to the second modulation back to the original key. Horizontal alignment is offset slightly to make overlapping results easier to distinguish. Error bars indicate estimated standard error.

TABLE 2. Post hoc *Multiple Comparisons by Type for Experiment 1.*

Type	Duration (seconds)	Significant Comparisons
I	0	3, 21
	3	All
	7.5	3, 21
	13.5	3, 21
	21	All
II	0	7.5, 13.5
	3	13.5
	7.5	All except 3
	13.5	All
	21	7.5, 13.5
III	0	3, 7.5
	3	All except 7.5
	7.5	All except 3
	13.5	3, 7.5
	21	3, 7.5

Note. Significant comparisons were determined by Tukey-Kramer tests.

the other hand, Type I and Type II responses diverged considerably for longer Section 2 durations. Given that the most salient difference between Type I and Type II sequences was the presence or lack of authentic cadences in the new key, the inclusion of cadences appeared to establish the new key more firmly, resulting in greater tension responses upon return to the original key.

Responses to Type I sequences with durations longer than 3 s had approximately neutral or positive values, pointing to a general increase in tension upon return to the original key as the duration of the new key increased. The relationship was not precisely linear though: the mean slopes for 13.5 s were slightly lower than for 7.5 s. This discrepancy might be explained by the fact that there was a very clear cadence at the end of the 7.5 s segment that essentially ended on a prolonged tonic ($I–I^6$), thus concluding Section 2 with a more anchored harmonic context and eliciting a stronger tension response to the return to the original key despite the shorter time span in the new key.

The mean values for Type II, on the other hand, were all negative, although the difference between 21 s and the control were not statistically significant. Furthermore, with the exception of the 21 s condition, the negative values increased as the Section 2 duration increased. It appears that the longer there was no cadence in the new key, the greater the release of tension when a cadence finally did occur—albeit in the original key. However, by 21 s, the memory of the original key had faded to a significant extent and the new key was fairly established despite the lack of a clear cadence.

The mean values for Type III sequences represent the clearest window into the time course of the memory decay of the original key. Tension increased linearly in

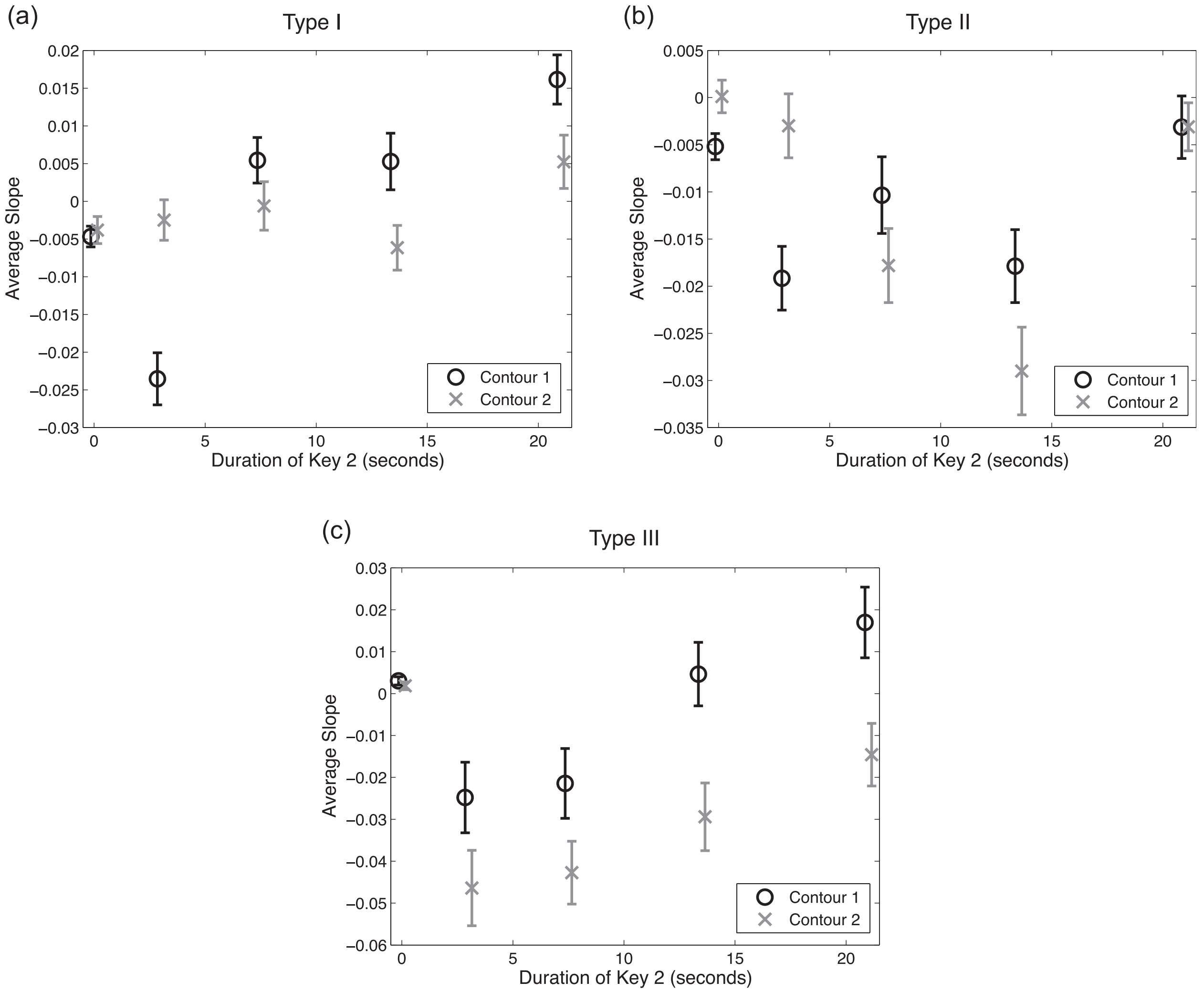

FIGURE 6. Comparison between opposing melodic contours for Experiment 1. (a) Type I (b) Type II (c) Type III. Horizontal alignment is offset slightly to make overlapping results easier to distinguish. Error bars indicate estimated standard error.

a very regular fashion as the duration of Section 2 increased. The mean values for the 21 s case were close to zero, similar to Type II. These results suggest that in the absence of significant interference due to tonal context, memory decay of the original key is quite gradual.

Contour comparison. The differences in mean values between the two opposite contours for each sequence are shown in Figure 6. Paired *t*-tests (two-tailed) between the slope means of the two differing contours in each Duration × Type category resulted in significant or near significant differences at an alpha level of .01 (to correct for multiple comparisons within type) consistently in the 3 s case. Unlike the sequences with modulations, the sequences with no Section 2 (0 s case) did not have exactly opposing contours; for Types I and II, both contours moved downwards, although to varying degrees. That likely explains why the 0 s slopes tended to be slightly negative on average. Most noteworthy was the marked contrast in responses to opposing contours in the 3 s condition for Types I and II. These differences leveled out with increasing duration, although in almost every case, the upward contours had greater slope values. This was seen most clearly for Type III, where the relative differences in slopes for the two-contour cases were almost identical, while the upward contours elicited consistently higher slope values than the downward contours. The results of all the *t*-tests are shown in Table 3.

Music training. As noted earlier, most participants had some degree of musical experience, although the

TABLE 3. *Contour Comparison t-test Results for Experiment 1.*

Sequence Type	Section 2 Duration	*t*	*p*
Type I	0	0.38	.70
	3	5.44	< .001*
	7.5	1.66	.10
	13.5	2.42	.02
	21	2.36	.02
Type II	0	2.39	.02
	3	3.95	< .001*
	7.5	1.51	.13
	13.5	2.25	.03
	21	0.009	.99
Type III	0	0.90	.37
	3	2.67	.009*
	7.5	2.38	.02
	13.5	4.08	< .001*
	21	3.94	< .001*

Note. df = 99 in all cases. **Significant results*

extent of training was highly variable. The effect of music background on responses was examined by dividing the subject pool into two categories: those having more and less music training. The "more" category consisted of subjects who had at least 7 years of training on their primary instrument, self-rated at least a 4 out of 5 in proficiency on their primary instrument, self-rated at least a 3 out of 5 on their overall music training, and had at least one semester of college-level music theory. All others were placed in the "less" music training category, resulting in 22 subjects in the "more" category and 28 subjects in the "less" category. The mean slopes for each Duration × Type case for the two categories are shown in Figure 7. There were no statistically significant differences at a .01 alpha level; however, the more positive values for subjects in the "less" music training category for both Types I and II sequences (in the latter case, these difference appear very consistent, if not statistically significant when comparing individual means independently) may be an indication that subjects with less music training were less sensitive to the presence of cadences. This is consistent with results from previous studies showing that musicians are more sensitive to harmonic structures than nonmusicians (Bigand & Parncutt, 1999; Bigand et al., 1996; Farbood, 2012).

Early versus late trials. It could be argued that given the repetition of the stimuli and the obvious structure of the sequences, there might have been a learning effect over the course of the experiment. Subjects might have started to anticipate the return to the original key and this building expectation could have affected the results. To examine this issue more closely, the mean slopes for only the first five trials of each type per subject were compared to the last five trials for each type. The early and late trials for each type were extracted strictly with regard to presentation order, which meant that some of the duration groups differed in size and the contour types were also unequally distributed (Figure 8). Nonetheless, *t*-tests indicate there were no significant differences; the only near-significant difference was the 21 s condition for Type I. This shows that learning effects had a minimal impact on the results. This is perhaps not surprising given that schematic expectations, informed by tonal context, facilitate processing and are independent of veridical expectations (Justus & Bharucha, 2001).

DISCUSSION

The results of Experiment 1 generally showed that decreases in tension upon a return to the original key were evident up to 13.5 s in a new key. The longest duration of 21 s elicited greater tension responses than all shorter durations within each type, indicating that the memory of the original tonal center had faded to a significant extent. The 21 s responses for Type I were considerably greater than the Type II and Type III responses, suggesting that the original key had been forgotten to a greater extent for Type I sequences. The 21 s responses for Types II and III were close to zero (very slightly negative), indicating that while the memory of the original key had faded, it was not completely absent. The fact that Type II and Type III lacked cadences in the new key area might explain this contrast. For the mid-range durations of 7.5 and 13.5 s, the pattern of responses varied depending on type and appeared to be influenced by cadences as well. Whether the cadences were in the new key area in the case of Type I, or only present in the V^7–I transition back to the original key for Type II, the presence of these cadences had a striking effect on tension judgments when compared to Type III sequences, which completely lacked functional harmony in both the original and new key sections.

Given these somewhat complex results, an unresolved question still remained: if 21 s in a new key is still not long enough to completely erase the memory of a previous key, how long is actually necessary for a true tonal "reset"? Experiment 2 was designed to provide additional data toward the goal of answering this question.

Experiment 2

The design of the stimuli for Experiment 2 was similar in basic structure to Experiment 1: Section 1 was intended

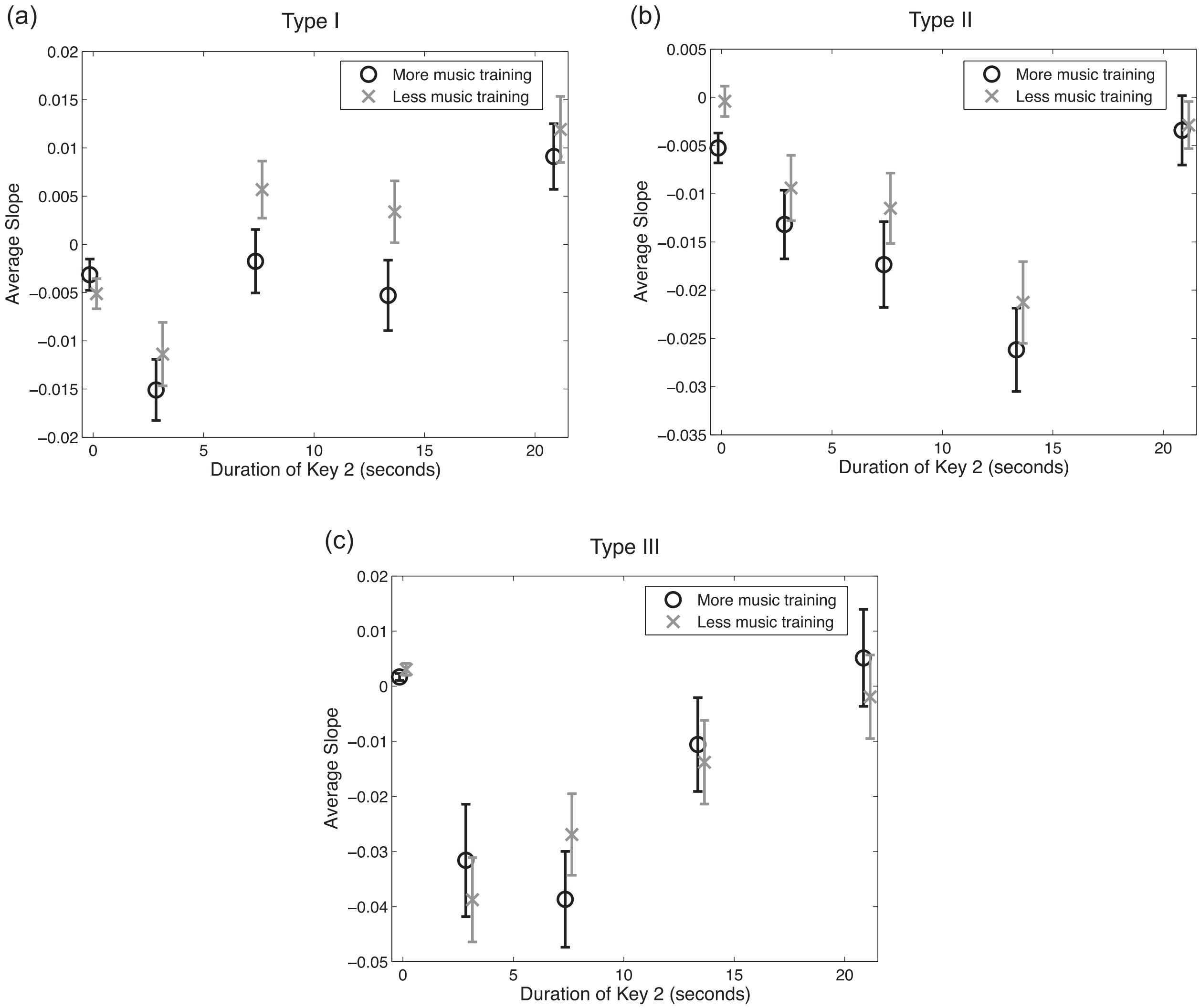

FIGURE 7. Comparison between subjects with more and less music training for Experiment 1. (a) Type I (b) Type II (c) Type III. Horizontal alignment is offset slightly to make overlapping results easier to distinguish. Error bars indicate estimated standard error.

to firmly establish tonic in a key area, the initial key modulated to another key in Section 2, and in Section 3 there was a return to the original key or a modulation to a third, completely new key. Since Type III sequences had yielded the most straightforward results in Experiment 1, the general style of the progressions for Experiment 2 were designed to be similar. In particular, all of the progressions had only a single repeated tonic chord in the new key area. The main difference between the new Experiment 2 progressions and the old Type III progressions was that Section 1 of the new sequences was more similar to Types I and II; instead of repeating a single tonic chord, there was a clear progression and cadence in the starting key. This was to ensure the original key was established as definitively as possible in addition to providing a way to alleviate the monotony of the stimuli. Unlike Experiment 1, there was only a single type of harmonic context for chord progressions (as opposed to three), and Section 2 durations were now considerably longer.

METHOD

Participants. A total of 31 participants took part in Experiment 2, most of whom were undergraduate and graduate students at New York University, mean age 24.60 years ($SD = 4.66$), 21 male. There was no overlap with the participant pool from Experiment 1. Subjects had an average of 7.79 years of formal training on a primary musical instrument ($SD = 3.77$) and an average self-rank in instrumental skill level of 3.77 ($SD = 1.05$)

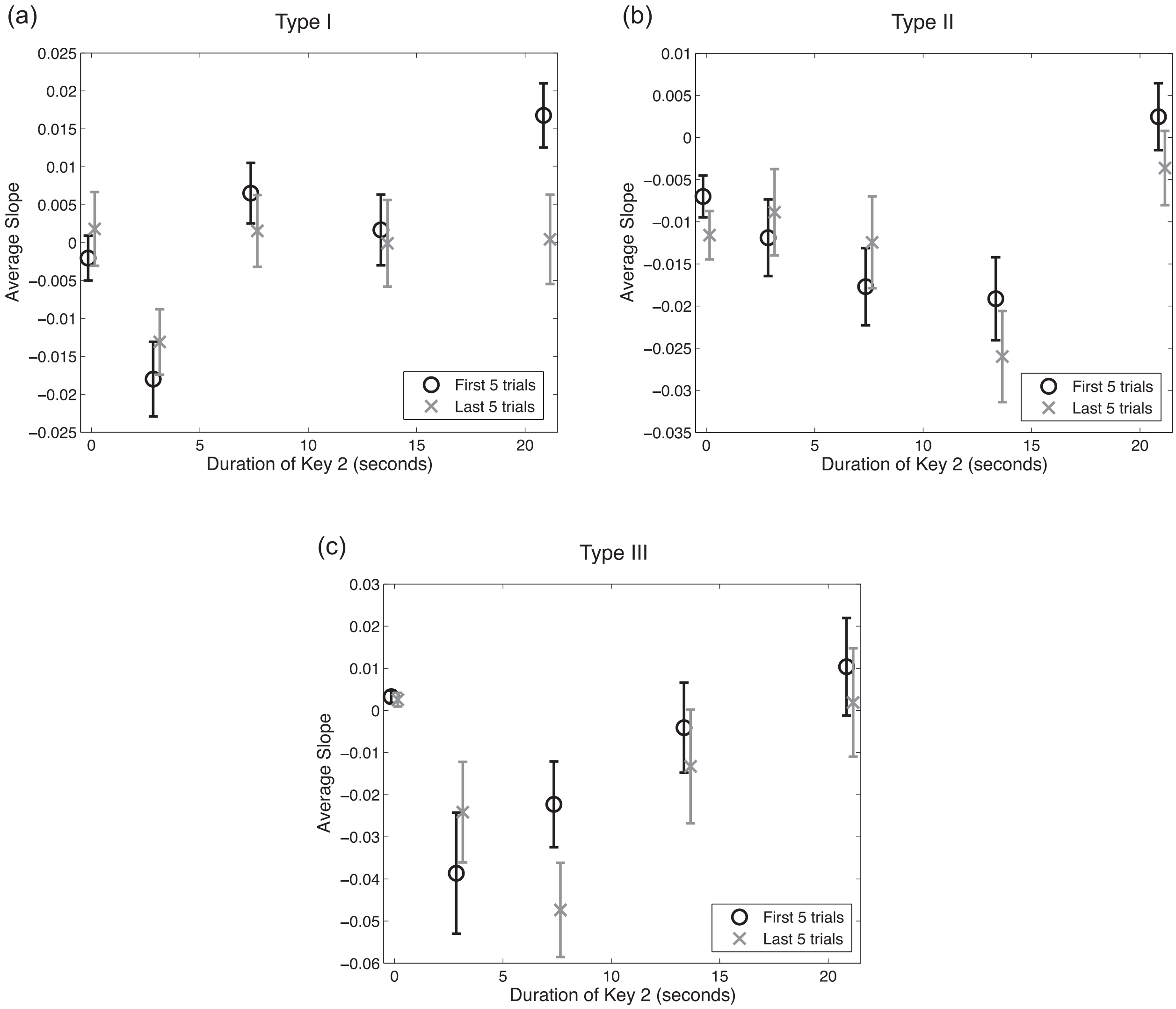

FIGURE 8. Comparison between the first five and last five trials of each type for Experiment 1. (a) Type I (b) Type II (c) Type III. Horizontal alignment is offset slightly to make overlapping results easier to distinguish. Error bars indicate estimated standard error.

on a scale of 1-5. Average number of semesters of college-level music theory training was 3.63 ($SD = 2.34$) and the mean overall self-ranked music training level was 3.86 ($SD = 0.91$), where 0 = *no training* and 5 = *professional-level training*. Five subjects reported having absolute pitch and one reported having partial absolute pitch.

Stimulus materials. The stimuli were rendered in the same way as Experiment 1. The durations of Section 2 were 0, 2, 10, 20, 30, and 45 s. All of the progressions prior to transposition are shown in Figure 9. Sequences 1-6 featured modulations to different keys and included two sequences (3 and 5) in minor keys. The addition of minor sequences was intended to add some harmonic variety to the listening experience. Sequences 1-6 were "closed" sequences (starting and ending in the same key). They were also designed to return to the original tonic without changes in the highest (melodic) pitch. This allowed for fewer stimuli since two harmonically identical sequences with opposite pitch contours were not necessary to balance the effect of pitch height. Although there may have been some residual effects of contour due to inner voice or bass line movement, the effect was significantly mitigated by holding the "soprano" line constant.

Sequences 7 and 8 (two additional versions of Sequence 6) were used to examine comparisons between

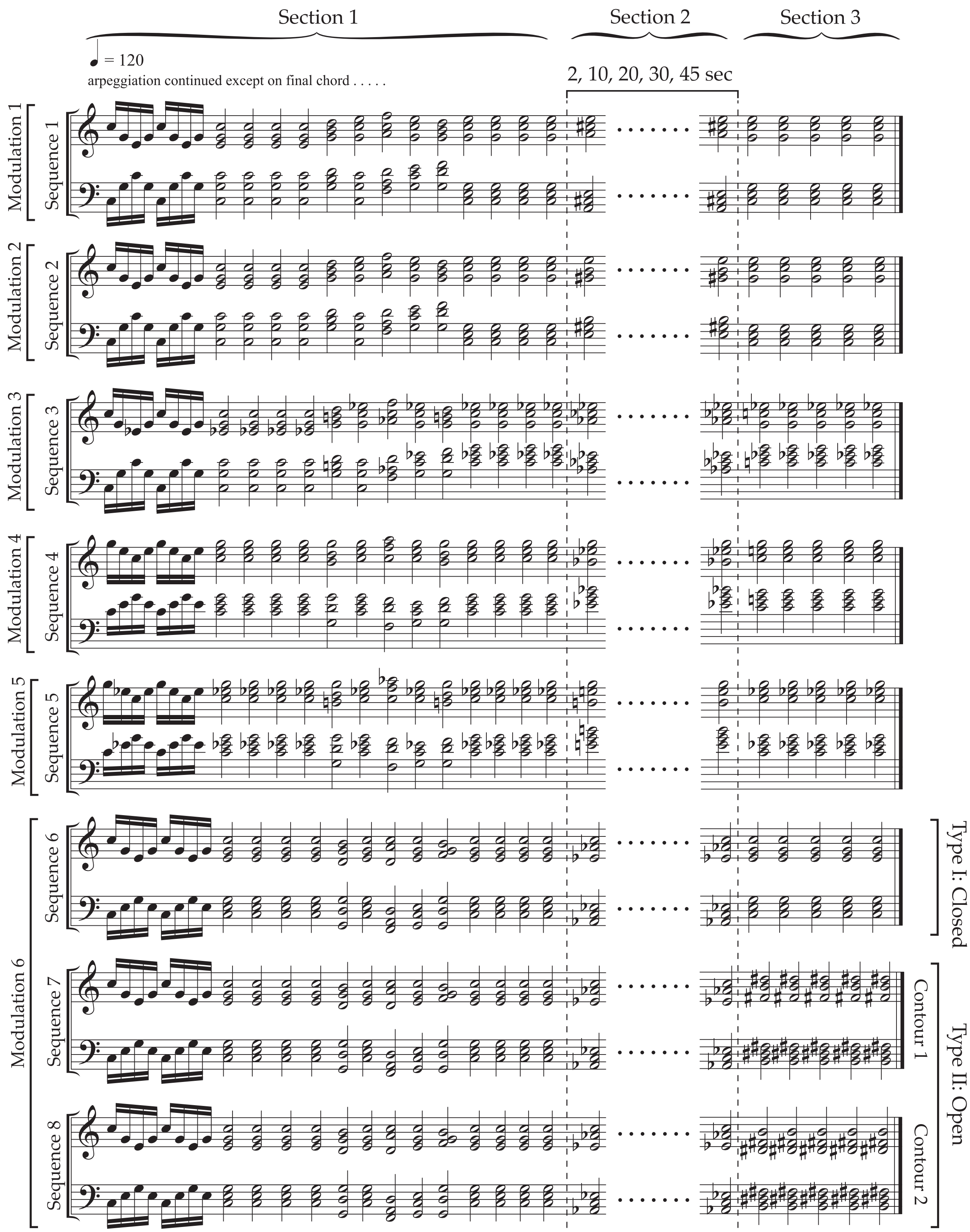

FIGURE 9. The harmonic sequences composed for Experiment 2.

responses to open and closed progressions. Both Sequences 7 and 8 modulated to the same key as Sequence 6 in Section 2, but instead of returning to the original key in Section 3, they modulated to a completely new key; thus, they were "open" progressions in the sense that they did not return to the original tonic. Sequences 7 and 8 were identical except one moved up in melodic contour when modulating to the third key area and the other moved down. Unlike Sequences 1-6, two harmonically identical sequences (Sequences 7 and 8) were necessary to create the open condition. This was due to the impossibility of finding a third key that was at least three steps on the circle of fifths distant from both the first and second keys in which a common tone was shared in the upper voice between tonic chords in Sections 2 and 3.

Procedure. The procedure for Experiment 2 was identical to Experiment 1. Likewise, all sequences were transposed pseudorandomly, ranging from 11 semitones down from the original sequence to 11 semitones up. All transpositions had to be at least three steps on the circle of fifths away from the key of the previous sequence, alternating between transpositions up and down. There were four practice trials followed by 45 recorded trials presented in random order. In summary, the trials included three types of sequences:

1) *Nonmodulating* - five sequences with no modulations from either Sections 1 to 2 or Sections 2 to 3, all identical except for the duration of Section 2.
2) *Closed* - 30 sequences modulating from Section 1 to Section 2 and then back to the original key in Section 3; six sequences (Sequences 1-6) $\times$ five Section 2 durations.
3) *Open* - 10 sequences modulating from Section 1 to Section 2 and then to a third key in Section 3; two sequences (Sequences 7-8) $\times$ five Section 2 durations.

RESULTS

As in Experiment 1, all data were normalized within subject as *z*-scores and slope values were calculated by looking at the sections of the continuous responses immediately following modulations. Since the modulations to Section 3 were quick transitions without a pivot chord (same as the transitions in Type III sequences from Experiment 1), the analyzed sections consisted of 1.5 s of continuous response data following each key change.

As expected, the average slope values following the initial modulation from Section 1 to Section 2 for both the open and closed sequences were large and positive (and nearly identical), while the corresponding values

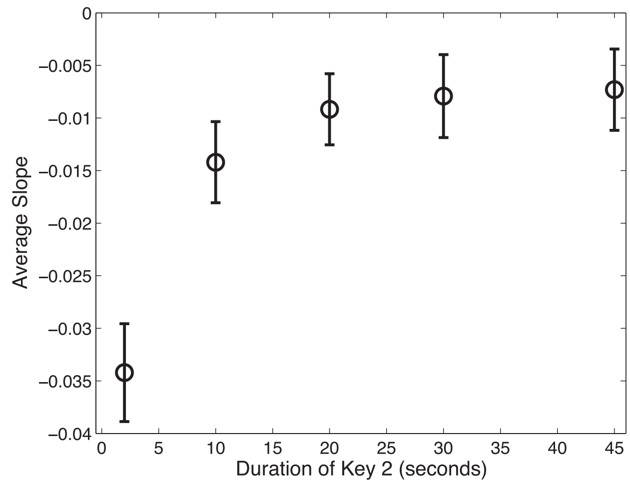

FIGURE 10. Experiment 2 results. Mean tension slopes in response to the second modulation back to the original key across all closed modulations. Error bars indicate estimated standard error.

for the nonmodulating sequences were markedly smaller: open $M = 0.06$, $SD = 0.05$; closed $M = 0.06$, $SD = 0.05$; nonmodulating $M = 0.004$, $SD = 0.02$. Responses to modulations from Section 2 to Section 3 were examined next. The mean slopes for each Section 2 duration across all closed trials are shown in Figure 10. The results showed that the values increased as the duration of Section 2 increased, with the greatest jump between 2 s and 10 s. A one-way, repeated-measures ANOVA revealed a significant effect of Section 2 duration, $F(1.54, 46.15) = 3.76$, $p = .04$, $\eta_p^2 = .11$, $\eta_G^2 = .06$. Post hoc Tukey-Kramer tests showed significant pairwise comparisons between only the 2 s condition and all others (10-45 s), but not between any of the longer durations.

Closed versus open progressions. When examining responses to modulations from Sections 2 to 3 in closed versus open progressions, only trials featuring Sequences 6-8 were compared directly, since they were identical until the start of Section 3. For analysis purposes, the slope values for Sequences 7 and 8 were merged by averaging. This was necessary to balance the number of data points between the open and closed groups; this also served to combine the data corresponding to the differing contours of the open stimuli. The mean tension responses for both types of progressions are juxtaposed in Figure 11, and a direct comparison of the slope values at the beginning of Section 3 is shown in Figure 12. There was a consistent pattern of responses where the slopes for trials modulating to a completely new key (open progressions) were larger

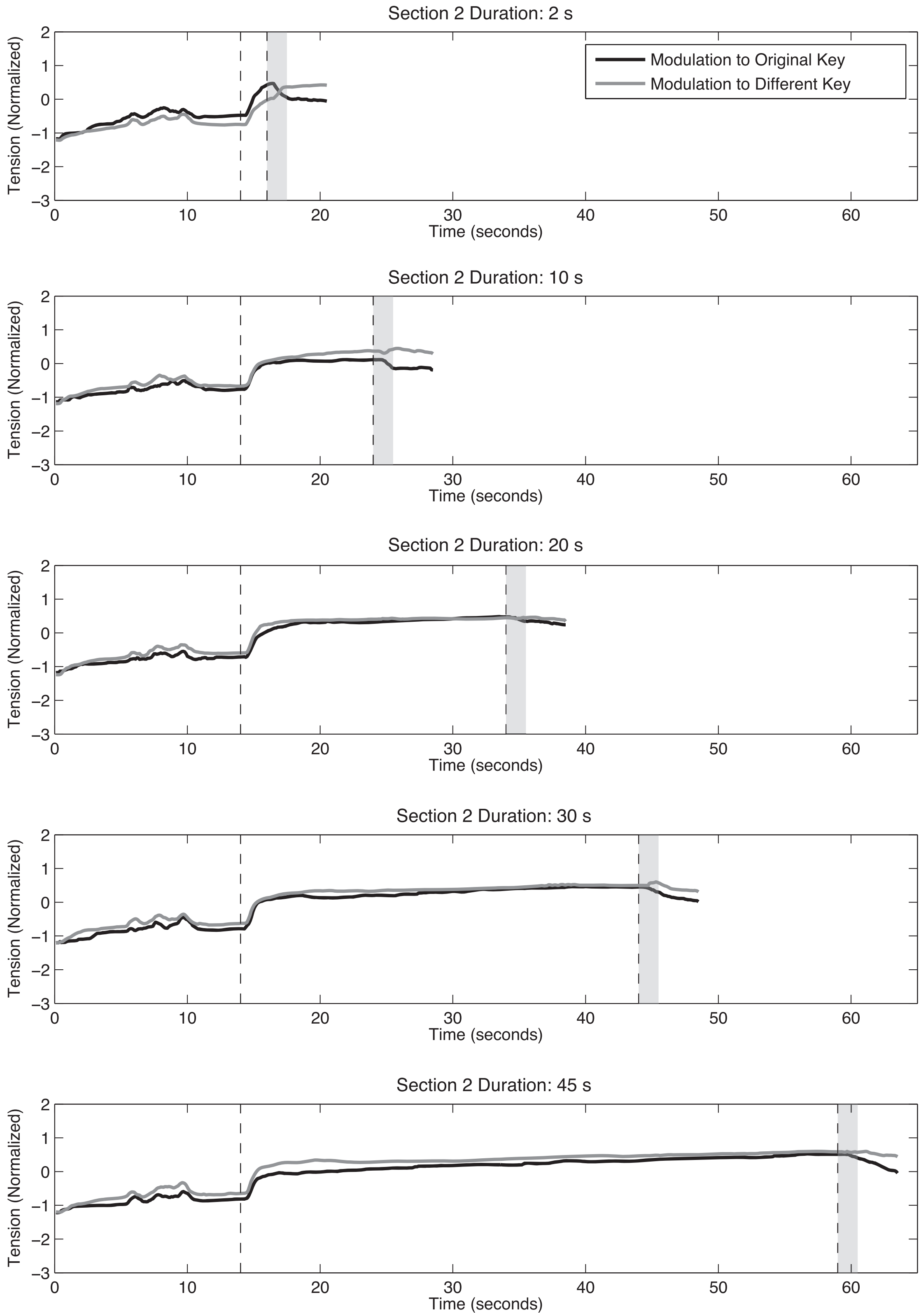

FIGURE 11. Mean tension responses for closed versus open sequences from Experiment 2.

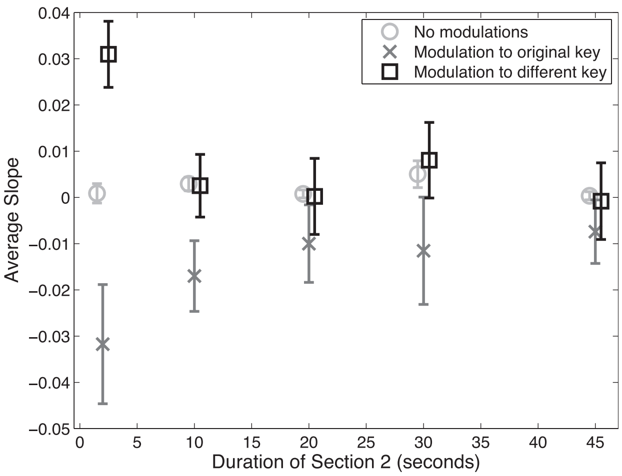

FIGURE 12. A comparison between Experiment 2 results for open, closed, and nonmodulating sequences. Horizontal alignment is offset slightly to make overlapping results easier to distinguish. Error bars indicate estimated standard error.

TABLE 4. *T-test results for comparison between open and closed conditions in Experiment 2.*

Section 2 Duration	*t*	*p*
2	3.97	< .01*
10	3.07	< .01*
20	1.23	.23
30	2.69	.01*
45	0.64	.53

*Note. *Statistically significant; df* = 30 in all cases.

than those modulating back to the original key (closed progressions). Two-tailed *t*-tests comparing values for open versus closed conditions for each Section 2 duration revealed statistically significant differences (at an alpha level of 0.1 to account for multiple comparisons) for the 2, 10, and 30 s conditions (Table 4).

DISCUSSION

The results of Experiment 2 showed that slope values for responses to closed progressions increased as the duration of Section 2 increased, with the greatest difference being between 2 s and 10 s. Despite the monotonic increase in slope corresponding to increase in the Section 2 durations, the differences between longer durations (10-45 s) were not statistically significant. Although all of the mean slope values were negative, indicating decreases in tension, this may be misleading. The nature of the stimuli could have affected the responses, since the long periods of repetitive harmony may have elicited a strong anticipation of change. There is a gradual upward trajectory in the tension responses over the course of Section 2 (visually evident in Figure 11). This trend is not seen in responses to Type III progressions from the Experiment 1. The cause of the trend in Experiment 2 could have been due to the monotony and predictability of the stimuli resulting from the longer Section 2 durations as well as the uniformity of the stimuli across trials. Post-test interviews revealed that many subjects felt some degree of relief when there was a change after hearing the same chord repeated for an extended period of time. It seems likely that a sense of released tension caused by *any* change interfered with listeners' perception of harmonic tension change. This conclusion is further supported by responses to the open sequences. As shown in the response profiles, the modulation to the third key area in Section 3 did not elicit the same type of strong reaction as the initial modulation to Section 2.

Regardless of this issue, the results of the comparison between open and closed variants of a progression (Sequence 6 vs. Sequences 7-8) indicated a divergence in responses at the point of the second key change. Modulations to an entirely new key elicited responses with more positive slopes, whereas modulations back to the original key elicited responses with more negative values. The divergence was in the same direction in all cases and statistically significant for in the 2, 10, and 30 s condition. All of the slope values corresponding to the closed condition were negative; all of the values corresponding to the open condition were positive except the 45 s case, which was close to zero. However, given the likely interference effects due to change anticipation, there was not enough evidence to conclude that the memory of the original key was retained beyond 20 s.

General Discussion

The results of this study offer a new perspective on how tonal centers are retained in working memory. Harmonic tension judgments following modulations from an established key to a new key and then back to the original key were evaluated by examining the tension slopes at those critical moments. The initial modulation to a new key elicited a spike in the tension as expected and the modulation back to the original key resulted in a distinct pattern of tension changes that varied depending on the duration of the new key area and the type of local harmonic progression in the new key. There were three harmonic context types explored in Experiment 1: Type I sequences, which contained stylistically expected chord progressions punctuated by authentic cadences in the new key, Type II sequences, which featured meandering harmonies that lacked cadences, and Type III sequences, which consisted of only a repeated tonic triad. Although the recollection of the original key faded with increasingly long time spans in a new key, some traces appeared to remain even after 21 s, the longest duration tested in Experiment 1. Experiment 2 explored longer time spans in a new key, up to 45 s, using progressions that combined elements of Type III sequences (the single repeated tonic in the new key area) with introductory material similar to Types I and II (clear cadence to establish initial key area). Experiment 2 also compared responses to closed versus open sequences, where closed sequences returned to the initial key after some time in a new key area while open sequences modulated to a third, completely new key. Although there were negative mean slopes in response to all conditions for closed sequences and positive slopes in all but the 45 s condition for the open sequences, there was not enough evidence to conclude that the memory of the original key persisted beyond 20 s.

TONAL CLOSURE

The findings for Experiment 1 also revealed that the recollection of the original key was influenced not only by the duration of the new key area, but by the harmonic context of the new key area as well. Although the memory of the original key faded increasingly with longer time spans in a new key, the decay was more marked for Type I sequences, which incorporated cadences in the new key. These results highlight the importance cadential closure has on the memory of a previous key, aligning with Schenker's (1906/1954) observation that "in general, a cadence in the new key has proved to be the most suitable means to fortify the new key and thus to make the modulations real and complete." Cadential closure is regarded as a foundational concept in tonal theory and has been discussed from both schematic and prolongational perspectives. This does not necessarily imply a dichotomy, however, as schemata and prolongation are not mutually exclusive concepts—cadential schemata play a role in prolongational theories as well as other theories of musical hierarchy. Nonetheless, from a psychological viewpoint, schemata tend to be discussed in terms of local structures, while prolongation, by definition, refers to structures that are influential over time.

Past studies have found that harmonic schemata are the highest predictor of strength of closure (Parncutt, 1995; Rosner & Narmour, 1992). These learned schemata are central to the perception of tonal style. The presence of the cadences in Type I sequences thus serve as powerful stylistic markers that affect tonal memory judgments (Cook, 1987; Spitzer, 1996). Spitzer (1996) remarked that "while our sense of tonal centricity might decay over time, our sensitivity to the 'character' of cadences is more stable. A gesture may sound conclusive in any key . . . a theory of long-range hearing must therefore be predicated on the negotiation of chains of structural markers, rather than on a memory for tonal centre." Psychological applications of prolongational theories must therefore impose perceptual limits on how listeners recall tonal centers after modulations. The importance of local schemata also suggests that simply adding a decay component to higher-level branches of Lerdahl's prolongational hierarchy, as suggested by Farbood (2010), would not completely account for the tension responses to the Type I sequences; an additional interference component that takes into account cadences in a new key is most likely necessary (cf. Altmann & Schunn, 2012; Berman, Jonides, & Lewis, 2009). This is supported by the findings of Collins, Tillmann, Barrett, Delbé, and Janata (2014), who derived a model of tonal expectation from behavioral data that included periodicity pitch distributions, chroma vectors, and activations of tonal space as variables. While their model successfully explained the patterns in the data, they found that adding a cadential closure component significantly increased the explanatory power of the model.

MODULATION AS SENSORY AND COGNITIVE PROCESSES

In addition to the presence of cadences, another possible form of interference is modulation distance. Krumhansl and Kessler (1982) used a probe-tone paradigm to test fitness of all chromatic pitches at various points in harmonic sequences that modulated to near or distant keys. They concluded that modulation between close keys was more quickly actualized than between relatively distant keys and that the persistence of the original key was stronger in the case of close keys. They observed that difficulty or ease of modulation was further affected by how the modulating sequence was constructed—the length and the path of modulation were both factors. The stimuli used in their experiment were relatively brief (approximately 5 s), and the design of the experiment was not intended to implicate a precise timescale for the integration process, therefore it is difficult to draw conclusions about how key distance affects memory over longer time spans. Leman (2000) proposed that sensory rather than cognitive factors might account for Krumhansl and Kessler's observations, demonstrating that a model based on echoic memory images of periodicity pitch could account for their results. However, he noted that although his model appears to model degree of fitness for a probe tone in a tonal context successfully, a schema-based model is still required for actual recognition of a tonal center. This observation is echoed in the results of the modeling study by Collins et al. (2014), mentioned above in the context of cadential closure. Collins et al.'s model used both sensory and cognitive variables to model reaction time data from several harmonic priming studies. While both cognitive and sensory components were significant to the model, the varying weights of the variables suggested that cognitive influences contributed more to tonal expectation than sensory aspects.

Although it seems possible that the distance of a modulation affects key memory (whether due to sensory or cognitive influences), it was not evident in the current study. In Experiment 1, a mid-distant modulation type (three steps away on the descending circle of fifths) was intentionally chosen for Type I and II sequences in order to serve as an "average" modulation. Type III sequences on the other hand were designed to have no overlap in pitches between the tonic of the initial key and the tonic of the original key and had no functional transitions between sections. Thus, Type I and II sequences had significant pitch class overlap between the original and new keys while Type III sequences had none. Despite these differences, the tension responses in the 21 s condition for Type II were highly similar to Type III and quite different from Type I, suggesting that the presence of cadences was a more significant interference factor than key distance across longer time spans.

Remarking on cognitive functions in tonal perception, Krumhansl (1983) noted that "although the sensory encoding of frequencies, amplitudes, and durations of the tones is a necessary stage, the information is presumably recoded, organized, and stored in memory in a form that may be quite unlike early sensory codes." Tonal induction appears to be part of this recoding and reorganization process and has been directly observed in functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) experiments on modulation perception (Janata et al., 2002; Koelsch et al., 2003). Janata et al.'s (2002) fMRI study presented listeners with a continuous melody that cycled through all 24 major and minor keys. They observed that the areas of the brain that tracked key changes corresponded to cortical regions associated with cognitive, affective, and mnemonic processing. They concluded that the structure of tonality was maintained in these regions as a dynamic topography. Koelsch et al. (2003) observed that modulations elicited a unique brain response (a slow negativity apparent around 500-1500 ms) not found in responses to other violations of musical regularity such as tone clusters, Neapolitan chords, and secondary dominants. They hypothesized that this response reflected cognitive operations encompassing the integration process that accompanies a change of key.

These neuroimaging studies, as well as numerous behavioral studies (probe-tone and priming experiments in particular), have explored mental representations of tonality from multiple angles. However, given the relative dearth of research focusing on real-time aspects of tonal induction, the current work offers an additional perspective on the time course of this cognitive integration process. These results suggest that although a modulation may happen fairly quickly, it takes a considerably longer period of time before the process is actually complete.

Author Note

Correspondence concerning this article should be addressed to Morwaread Farbood, Department of Music and Performing Arts Professions, 35 W. 4th St., Suite 1077, New York, NY 10012. E-mail: mfarbood@nyu.edu

References

Altmann, E. M., & Schunn, C. D. (2012). Decay versus interference: A new look at an old interaction. *Psychological Science*, *23*, 1435-1437.

Auhagen, W., & Vos, P. G. (2000). Experimental methods in tonality induction research: A review. *Music Perception*, *17*, 417-436.

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 317-333. DOI: 10.1037/a0014873

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

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.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception and Psychophysics*, *41*, 519-524.

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research*, *62*, 237-224.

BIGAND, E., & PINEAU, M. (1997). Global context effects on musical expectancy. *Perception and Psychophysics*, *59*, 1098-1107.

BIGAND, E., PARNCUTT, R., & LERDAHL, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and music training. *Perception and Psychophysics*, *58*, 125-141.

BIGAND, E., TILLMANN, B., POULIN-CHARRONNAT, B., & MANDERLIER, D. (2005). Repetition priming: Is music special? *The Quarterly Journal of Experimental Psychology Section A, Human Experimental Psychology*, *58*, 1347-1375. DOI:10.1080/02724980443000601

BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433-436.

COLLINS, T., TILLMANN, B., BARRETT, F. S., DELBÉ, C., & JANATA, P. (2014). A combined model of sensory and cognitive representations underlying tonal expectations in music: From audio signals to behavior. *Psychological Review*, *121*, 33-65. DOI: 10.1037/a0034695

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

CRAIK, F. I. M., & LOCKHART, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.

CUDDY, L. L., & THOMPSON, W. F. (1992). Asymmetry of perceived key movement in chorale sequence: Converging evidence from a probe-tone analysis. *Psychological Research*, *54*, 51-59.

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

FARBOOD, M. (2010). Working memory and the perception of hierarhical tonal structures. In S. M. Demorest, S. J. Morrison, & P. S. Campbell (Eds.), *Proceedings of the 11th International Conference of Music Perception and Cognition* (pp. 119–222). Seattle, WA: ICMPC.

FARBOOD, M. M. (2012). A parametric, temporal model of musical tension. *Music Perception*, *29*, 387-428. DOI:10.1525/mp.2012.29.4.387

GRANOT, R. Y., & EITAN, Z. (2011). Tension and dynamic auditory parameters. *Music Perception*, *28*, 219-246.

GRANOT, R. Y., & JACOBY, N. (2011). Musically puzzling I: Sensitivity to overall structure in the sonata form? *Musicae Scientiae*, *15*, 365-386. DOI:10.1177/1029864911409508

JANATA, P., BIRK, J. L., VAN HORN, J. D., LEMAN, M., TILLMANN, B., & BHARUCHA, J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, *298*, 2167-2170.

JUSTUS, T. C., & BHARUCHA, J. J. (2001). Modularity in musical processing: The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 1000-1011.

KARNO, M., & KONEČNI, V. J. (1992). The effects of structural interventions in the first movement of Mozart's Symphony in G-Minor, K. 550, on aesthetic preference. *Music Perception*, *10*, 63-72.

KLEINER, M., BRAINARD, D., & PELLI, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, ECVP Abstract Supplement.

KOELSCH, S., GUNTER, T., SCHRÖGER, E., & FRIEDERICI, A. D. (2003). Processing tonal modulations: An ERP study. *Journal of Cognitive Neuroscience*, *15*, 1149-1159.

KRUMHANSL, C. L. (1983). Perceptual structures for tonal music. *Music Perception*, *1*, 28-62.

KRUMHANSL, C. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, *13*, 401-432.

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.

LALITTE, P., & BIGAND, E. (2006). Music in the moment? Revisiting the effect of large scale structures. *Perceptual and Motor Skills*, *103*, 811-828.

LEMAN, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*, *17*, 481-509.

LERDAHL, F. (1996). Calculating tonal tension. *Music Perception*, *13*, 319-363.

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.

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

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

NIELSEN, F. (1987). Musical "tension" and related concepts. *The Semiotic Web '86: An International Yearbook*, 491-514.

PARNCUTT, R. (1995). Perceptual modeling of the finality of harmonic cadences. *Systematische Musikwissenschaft*, *3*, 71-85.

PELLI, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437-442.

ROSNER, B. S., & NARMOUR, E. (1992). Harmonic closure: Music theory and perception. *Music Perception*, *9*, 383-411.

SCHENKER, H. (1906/1954). *Harmony.* (E. M. Borgese, Trans., O. Jones, Ed.). Cambridge, MA: MIT Press.

SPITZER, M. (1996). The retransition as sign: Listener-orientated to tonal closure in sonata-form movements. *Journal of the Royal Musical Association*, *121*, 11-45.

THOMPSON, W. F., & CUDDY, L. L. (1992). Perceived key movement in four-voice harmony and single voices. *Music Perception*, *9*, 427–238.

Tillmann, B., & Bigand, E. (1996). Does formal musical structure affect perception of musical expressiveness? *Psychology of Music*, *24*, 3–17. DOI:10.1177/0305735696241002

Tillmann, B., & Bigand, E. (2001). Global context effect in normal and scrambled musical sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 1185-1196.

Tillmann, B., & Bigand, E. (2004). The relative importance of local and global structures in music perception. *Journal of Aesthetics and Art Criticism*, *62*, 211-222.

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

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception*, *16*, 99-117.

Tillmann, B., Janata, P., Birk, J., & Bharucha, J. J. (2008). Tonal centers and expectancy: Facilitation or inhibition of chords at the top of the harmonic hierarchy? *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1031-1043. DOI:10.1037/0096-1523.34.4.1031

Woolhouse, M., Cross, I., & Horton, T. (2016). Perception of nonadjacent tonic-key relationships. *Psychology of Music*, *44*, 802-815. DOI: 10.1177/0305735615593409