Real-Time Modulation Perception in Western Classical Music

Brendon Mizener[1] & W. Jay Dowling[1]

[1] University of Texas at Dallas

Author Note

Music Perception and Cognition Lab, Dr. W. Jay Dowling, Director. School of Brain and Behavioral Sciences, University of Texas at Dallas.

Correspondence concerning this article should be addressed to Brendon Mizener, 800 W. Campbell Rd., GR41, Richardson, TX 75080. E-mail: bmizener@utdallas.edu

Abstract

Music listening is an incredibly complex process. The task of listening involves an auditory scene analysis in which the listener processes a huge quantity of auditory information every second. Those judgments relate to melody, timing, and key or tonic region. Here we examine whether the process of tracking key region is independent of the process of tracking surface cues, and what surface cues may influence that process. To this end, highly-trained, moderately-trained, and untrained listeners listened to excerpts from string quartets, quintets, and sextets from the classical and romantic eras and responded when they heard a modulation. Each excerpt featured either a pivot chord modulation, a direct modulation, a common tone modulation, or no modulation. Listeners performed above chance across all training levels and modulation conditions, with musical features including modulation type, time, and mode change, as well as participant training level affecting overall accuracy. We also present an exploratory PCA as a means of establishing future directions. The results of the PCA suggest that harmonic language and phrasing are both significant factors in guiding modulation perception, both of which merit further investigation.

Keywords: Modulation, Tonic Area, Music Training, Reaction Time

Word count: 11161

Real-Time Modulation Perception in Western Classical Music

Listener understanding of pitch center, or tonic, is fundamental to understanding musical structures from the smallest musical motif to the large-scale structures that define musical form. The tonic in a musical phrase is the pitch that represents the foundation of that tonal hierarchy (Krumhansl & Keil, 1982; Krumhansl & Shepard, 1979; Lerdahl & Jackendoff, 1983). Listeners become aware of the tonic in a given composition through a number of musical and perceptual cues related to melody and melodic expectancy (Brown, 1988; Dowling, 1986; Lerdahl & Jackendoff, 1983) and harmony and phrase structure (Huron & Parncutt, 1993; Krumhansl & Kessler, 1982). Additionally, it has been well established that music listeners implicitly understand the concept of a key, or set of notes, and its tonic (Dowling, 1978; Krumhansl & Keil, 1982; Krumhansl & Shepard, 1979), but much less work has been done on understanding our perception of the motion between different tonic centers within a single composition. This process of moving between key centers is called *modulation*.

Modulation allows the composer to use a wider palette for musical expression. Western classical music theorists have defined several types of modulation, each of which is characterized by its own parameters and musical theoretical constructs. One question that arises from these concepts is whether or not the listener is perceptually aware of the composer's intent: does the listener perceive the progression from key center to key center with no more information than can be grasped by listening? Underlying this question is the question of how humans percieve tonality.

One model of tonal perception in humans comes from studies using the probe-tone method (Krumhansl & Shepard, 1979), which has often been replicated. This method involves participants rating the goodness of fit of a given note in a musical context, with the probe tone played either immediately after the music (Krumhansl & Shepard, 1979; Vuvan, Prince, & Schmuckler, 2011) or concurrently with the music (Raman & Dowling, 2016; Toiviainen & Krumhansl, 2003). The first use of this method was to rate how well each note in the chromatic scale completed an ascending scale (Krumhansl & Shepard, 1979). The results of this study established the model of the tonal hierarchy for major and natural minor keys. Vuvan, Prince, & Schmuckler (2011) extended this work to map the melodic and harmonic forms of minor. Toiviainen & Snyder (2003) further extended the probe-tone model to track listener perception of key in a dynamic context. This allowed the creation of a model that showed how the perception of tonality changed over time, beyond a static context. Further applications appear in Cuddy & Thompson (1992), who tracked modulations specifically and found that both trained and untrained listeners were able to judge the distance chord progressions modulated from the original tonic with a high degree of consistency. Additionally, Raman & Dowling (2017) performed an interesting cross-cultural study that showed effects of both training and enculturation in western and Indian listeners when tracking modulations in South Indian Classical (Carnātic) music.

Krumhansl and Kessler (1982) used an extension of the this to model perceived tonal organization by having listeners rate probe tones iteratively with each successive addition of a chord in a chord progression. They found both a recency effect for key information and a distance effect that resulted in cognitive lag for assimilation. The further a progression modulated, the longer it took for participants to process the modulation and internalize the new key. Krumhansl (1990) combined the data from Krumhansl and Shepard (1979) and those from Krumhansl and Kessler (1982) to create map of tonalities, projected on the surface of a torus, modeling the apparent cognitive distance between key centers.

Despite what some would describe as a functional deficit in the probe-tone method (Butler, 1989, 1990), it has been found to be robust, thanks to the replication and converging evidence presented in studies enumerated above. There are other models, including the rare intervals hypothesis (Butler, 1989; see also Browne, 1990, and Brown & Butler, 1981) and Huron and Parncutt's (1993) model. The latter incorporates short-term echoic memory, in an effort to address a concept that may seem obvious to music listeners, namely that: "both structural and functional factors" (Huron & Parncutt, 1993) contribute to the perception of tonality.

We would suggest that all of these mechanisms explain part of the apparatus for tonality induction. Given the multifaceted nature of music (melody, harmony, rhythm, timbre, and their interactions) and its presentation to listeners, it makes sense that listeners would need multiple deductive paradigms to figure out what is going on in the navigation of tonal space. The system likely incorporates all of this information simultaneously to arrive at an answer.

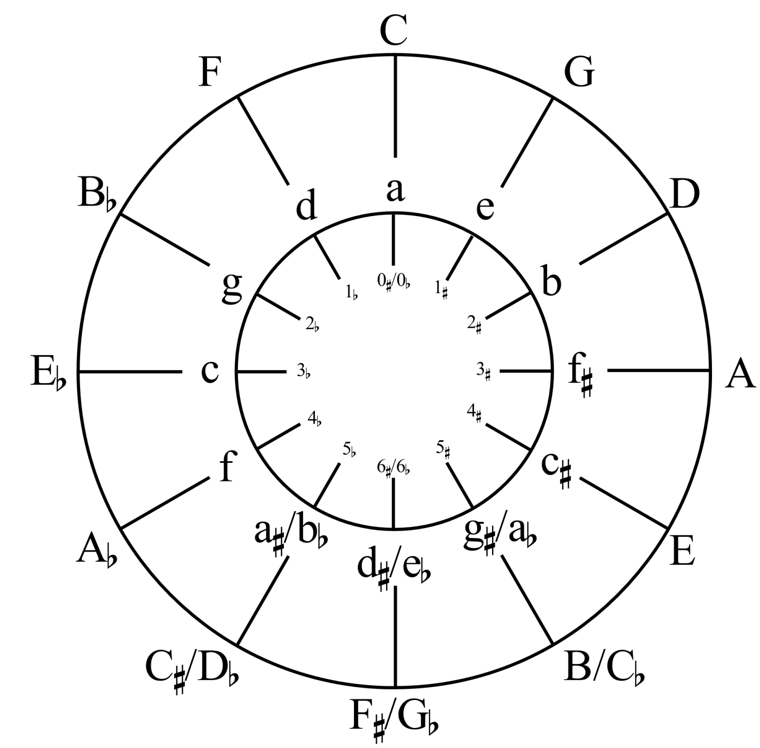
We attempt here to address some of the questions that arise from the probe-tone method. Specifically, because it depends fundamentally on a listener's rating of dissonance between a given chromatic pitch and the harmonic and melodic cues at any point during a piece, it doesn't address a whether or not listeners track key information, or, more relevant for the current study, information about the relationship between keys. One apparent deficit is seen in the dynamic map that is a product of Toiviainen & Krumhansl (2003). It appears to track chord progressions by way of micro-scale judgments of perceived dissonance, rather than the tonic area per se. Defenders of the probe-tone model argue that this is an oversimplification, and that the sum of the overall judgments made by listeners in the probe-tone model pinpoints the tonic area via correlations of patterns of judgments. That study, however, is one of only a few that uses naturalistic stimuli, which are used for the present study and should be a goal of future research on this topic.

Modulation as a Music-Theoretical Construct

Many of our ideas about music perception are based on a tradition in music theory that begins with Rameau's *Treatise on Harmony* (1722/1971). Fundamental to current models of tonal music perception is the concept of a central or foundational pitch: the tonic (Krumhansl, 1990). This is the note against which other notes and groups of notes are compared to guide our understanding of phrase, tonality, and closure (cadences) (Sears, 2015). However, as the tonic is not present in every moment of western music, our understanding of tonic is also shaped by the harmonic structures that surround it (Butler, 1989). A key or tonal hierarchy is the group of notes that surround the tonic, which are both defined by and, conversely, define the tonic itself. The pattern of keys is often presented visually as the circle of Fifths (see Figure 1).

Figure 1.

The circle of fifths.



Each letter around the circumference of the concentric circles represents the tonic of a major (outer) and a minor (inner) key. Moving from one vertex to another constitutes a change in the tonal center of a key; a modulation. Moving to either of the adjacent vertices in the current ring or the corresponding vertices in the other (for C: G, F, a, d, and e) is considered to be a modulation to a near key; moving further than that is considered a distant modulation. Because the tonic of a key is determined both by hierarchy and by the intervallic pattern of the notes that surround it, each new key requires sharps (which raise a note by a semitone) or flats (which lower a note by a semitone), and in a few cases, both.

The number of sharps and flats that each key requires are listed in the center of the circle. The key of C has no sharps and no flats. The number of sharps and flats offers one measure of similarity between keys - adding one sharp or flat changes one note relative to the key adjacent. More sharps or flats results in greater difference between keys. A mode change is a change from major to minor or vice-versa. Moving the key up or down a fifth does not require a change in mode, but moving to the relative or parallel minor does.[1]

Types of Modulation

The three types of modulation considered for this study are: direct modulation, common tone modulation, and pivot chord modulation. There are other types of modulation codified in the western classical tradition, but we selected these three as maximally distinct. Each of these types of modulation features specific characteristic surface features that distinguish it. Our hope with this selection is to determine what surface features, if any, are effective cues in listener identification of change in key.

Pivot Chord Modulation. A pivot chord modulation occurs when a composer uses a chord that is common to, and has a similar function[2] in, two different keys, serving as a pivot to move from one key to another (Benjamin, Horvit, & Nelson, 2003). For purposes of this experiment, we selected only excerpts in which the pivot chord was common in both its root (note on which the chord is based) and its quality (major or minor) in the two keys. Figure 2 illustrates the progression described here. The target key for this type of modulation is often the dominant, with a vi chord as a pivot.[3] This creates a vi – II^7 – V

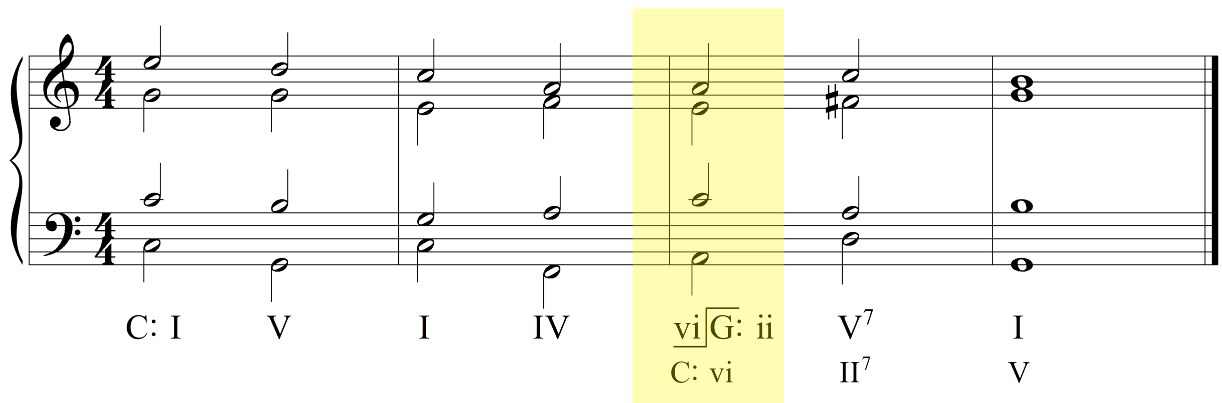
[1] However, as Krumhansl and Kessler (1982) notes, there are conflicting opinions on whether a shift to the relative minor counts as a key change at all, since there are no changes in the set of pitches that makes up the key.

[2] I.e. pre-dominant function, dominant function. This means that the chord in question serves a certain purpose within the phrase. Tonic is a place of rest, dominant is a place of tension, and pre-dominant sets up the dominant harmony.

[3] A note on typography: throughout this paper, chords are indicated in two ways: with roman numerals,

Figure 2.

Basic example of a pivot chord modulation.



Note. The progression is analyzed in C Major and G Major, with the roman numeral analysis in C after the modulation included for illustrative purposes. The highlighted chord is the pivot chord.

progression in the starting key that matches a ii – V^7 – I progression in the target key.[4] The example in Figure 3[5] shows a similar progression. The highlighted area indicates a I - V - vi^7 in the key of E♭, where the vi^7 serves pre-dominant function in both keys.

Direct Modulation. Though there is a technical distinction between direct and phrase modulation, we consider them functionally equivalent here. Phrase modulation occurs at a phrase boundary: a composer finishes a phrase in one key and begins the next phrase in a new key immediately after, with no transition material. Direct modulation occurs when an

with the number indicating the scale degree on which the chord is built, and using letters. Letters also indicate scale degrees. Notes or scale degrees on which unaltered triads (chords consisting of 3 notes, each an interval of a third from the one below it) form major chords are denoted by capital letters or roman numerals, and those on which unaltered triads form minor chords are denoted by lowercase. Diminished chords are indicated by a lowercase letter or numeral with a °. Hat signs (ˆ) indicate scale degrees, i.e. $\hat{7}$ is the seventh scale degree.

[4] Note that ii in a major key should be minor, so the II with a seventh added is an altered chord (II^7)

[5] All examples transcribed using Finale v. 25 for Windows.

Figure 3.

Excerpt featuring a pivot chord modulation.

Note. Haydn op. 2 No. 3, Mvt. 1, ms. 1 - 27. The large and small highlighted areas represent the modulating phrase and the pivot chord, respectively. Source: Haydn (1765/1845), transcription by Mizener.

abrupt modulation happens somewhere other than a phrase boundary (Benjamin et al., 2003). These modulations are often made to closely related keys, as defined earlier. However, because of the abrupt nature of the modulation, modulating to a close key is not a necessity, and many composers have used this technique to modulate much further.

Figure 4 is an excerpt from Beethoven's String Quartet in F major, op. 18 No 1, Mvt. 4: Allegro, in which the modulation moves from F major to d minor (tonic to relative minor). The initial I^{6}_{4} - V - I progression clearly establishes F major before moving back and forth between tonic and dominant harmony. Immediately on the downbeat of the eleventh measure, Beethoven switches suddenly to a d minor harmony by landing aggressively on the tonic of the V chord. The scale leading to the tonic, played in unison, uses the melodic minor scale by raising the B♭ to a B♮ and C♮ to a C♯ and landing on D on the downbeat of the twelfth measure. This dominant-tonic motion and the subsequent a melodic minor scale over the next two measures solidly establish the d minor harmony. The underscore in the figure indicates the sustained tonic harmony implied by the scale in its entirety.

Figure 4.

Excerpt featuring a direct modulation.

Note. Beethoven String Quartet No. 1 in F Major, op. 18, No. 1, mvt. 4, ms. 17-39. The highlighted area represents the modulation from F major to D minor. Although there are two measures highlighted, there is no 'transition material' moving between the two keys, and the modulation occurs directly on the downbeat of measure 11 of the excerpt. Source: Beethoven (1800/1937), transcription by Mizener.

Common Tone Modulation. A common tone modulation occurs when a composer uses a single sustained or repeated pitch or dyad (two pitches sounding simultaneously) to link two keys. This can occur either in the middle of a phrase or at a phrase boundary. The common tone is present in both keys but can serve either similar or different function in either key. Often, but not necessarily, the pitch is present in the tonic chord of both keys

(Benjamin et al., 2003). Because the only feature linking the two keys is a single pitch or dyad, the two keys need not be closely related. This modulation serves as an efficient way to connect distant keys. Figure 5 is an excerpt from Franz Schubert's String Quintet in C Major, D. 956, movement 1. The sixth measure of the excerpt starts in G, and the violins both sustain the tonic across the barline, where the second violin moves down to an E$\flat$, using the progression $\hat{3}$-$\sharp\hat{2}$-$\hat{2}$-$\hat{1}$ in E$\flat$ and $\hat{1}$-$\hat{7}$-$\flat\hat{7}$-$\flat\hat{6}$. This containes notes that fit into both keys and allow for a gradual transition to the new tonic. The E$\flat$ in the second violin is supported by the second cello on the first beat of the next measure and then repeated by the second cello on the second beat of the measure, firmly establishing the new key.

Figure 5.

Excerpt featuring a common tone modulation.

Note. Schubert's String Quintet in C Major, D. 956, Movement 1, measures 74 - 84. The large and small highlighted areas represent the modulating phrase and the common tone, respectively. The dashed bracket at the top of the excerpt represents the specific area that was played for the participant. Material outside of the bracket included for visual continuity. Source: Schubert (1828/1965), transcription by Mizener.

Present questions

The questions we consider for the present experiment are as follows: Do music listeners passively retain information on key region independent of topical, salient features of the music? To what extent does training affect the storage, processing, and access to that information, if it exists? What topical features influence our understanding of key regions and the movement between them? What is the balance between melodic and harmonic features contributing to that understanding?

To be clear, we do not claim that these musical elements are truly independent. we seek rather to investigate the level to which they are considered independently, or the level to which each is effective when making judgments within a musical context. For a long time, studies surrounding key area processing focus solely on either melodic material (Bartlett & Dowling, 1980; Dowling, 1986; Krumhansl & Keil, 1982) or harmonic material (Thompson & Cuddy, 1989). Only recently have researchers begun to use more naturalistic stimuli, such as excerpts or MIDI recreations in their research (Toiviainen & Krumhansl, 2003). Krumhansl (1983) argues convincingly that the various features of a musical piece are all interdependent and contribute to the processing of key area. It remains a question, however, what musical information takes priority in terms of salience during listening, and whether or not that information receives similar priority during processing. For the purpose of this study, we consider both melody and harmony to be topical, salient features. The information that would be "background" is not the harmonic motion per se, or even the chord qualities, but rather the relative implications, or functions of the chords in the context of the key. These chordal functions are readily understood upon harmonic analysis of a written score, but may not be understood by passive or even active listening. We therefore propose four hypotheses. First, that groups of participants who have greater levels of training, across all modulation types, will be more accurate. Should this hypothesis be supported, it would represent evidence that training is a significant factor in key recognition retention. Should it not be supported, and we find that instead, accuracy is independent of training, it would suggest that training is not necessary for key retention. This in turn suggests that key tracking is obtained through passive listening to music.

Secondly, across levels of training, the responses to the pivot chord modulation will be the least accurate, responses to the direct modulations will be the most accurate, and the responses to common tone modulations will fall in the middle. This hypothesis most directly addresses the question of topical salient musical features, and makes two specific predictions. Firstly, that different types of modulations, and the topical features of each, will inform listener perceptions to different degrees, and that the movement between the two keys will be recognized with different levels of accuracy. Secondly, it makes a specific prediction about the relative levels of accuracy that each of the specific modulation types will see. If participants were to respond to all of the modulation types with equal accuracy, it would suggest that topical features and key area recognition are independent of one another. This would support the idea that the perception of movement between key areas, and therefore the recognition of tonic area per se, are unaffected by topical features.

Next, key distance and mode change will be more accurate predictors of modulation perception. Excerpts that feature a mode change as part of the modulation should be more accurately recognized, likewise with greater key distance. This hypothesis, like hypothesis number two, also deals with topical features. Should these features be found to represent more accurate responses, it would suggest that those specific features guide perception to a greater extent than other types of musical features.

Finally, we hypothesize that trained listeners will respond faster to the modulations than untrained listeners. Hypothesis four, like hypothesis one, addresses the question of training, but from a different perspective. Faster response time in trained listeners would suggest that the key area processing becomes more efficient as a result of training.

Method

Participants and group assignments

The majority of participants for this study were adults selected from the UT Dallas undergraduate SONA pool. These students were compensated with credit towards their psychology research exposure requirements. Some participants were also recruited from the music department at Northwestern State University in Natchitoches, Louisiana (NSULA), and some professional musicians and music educators from around the region were recruited through direct personal correspondence. Participants who were not students at UT Dallas were not compensated.

Approximately equal numbers of male and female participants participated in the study (M = 92, F = 87, NB[6] = 1), but gender was not considered. Participants' mean age was 22.88 (SD = 5.49). Participants were excluded if they met any of the following criteria: exposure to or training in South Indian Classical (Carnātic) Music; absolute pitch; or a hearing disability such as deafness, tinnitus, or amusia. Participants were evenly divided into three groups (n = 60) based on years of formal music training. Participants with zero to two years of formal music training were assigned to the nonmusician/untrained category, those with three to 10 years of formal training were assigned to the moderately-trained category, and those with 10 years or more of training were assigned to the highly-trained musician category. The nonmusician group had a mean of 0.63 years of training (SD = 0.92), moderately-trained musicians had a mean of 5.53 years of training (SD = 1.75), and highly trained musicians had a mean of 16.07 years of training (SD = 7.75). Additionally, participants who had less than 10 years of training but had completed an Advanced Placement™ or college-level aural skills training course also qualified for the highly-trained musician category. There were only four participants who met those criteria, who had nine, eight, five, and five years of training, respectively. For purposes of this study, formal music

[6] Non-Binary

training was considered any time spent pursuing music in a formal setting, including instrumental or vocal large ensemble experience, small ensemble experience, and private lessons.

Stimuli

Stimuli were selected from the string quartet, quintet, and sextet works of Joseph Haydn (1732-1809), Roman Hofstetter (1742-1815), Wolfgang Amadeus Mozart (1756-1791), Ludwig van Beethoven (1770-1827), Franz Schubert (1797-1828), and Johannes Brahms (1833-1897). The string quartet idiom was selected to most effectively control the effects of timbre. A full list of excerpts and recordings is included in Appendix 1. Each of the excerpts were chosen to meet specific structural criteria. These criteria included length, the presence of a single modulation, surrounded on either side by a stable tonic area, and enough time to instantiate the intial key before the modulation occurred. The shortest stimulus was 21.05 s and the longest was 59.66 s, with a mean length was 28.98 s. Each stimulus included at least six seconds of stable establishment of tonic before the modulation occurred.

Each excerpt was ripped from its source CD using fre:ac version 1.0.32 (Kausch, 2018), an open-source audio converter. Stimuli were presented as .wav files to ensure the highest quality audio signal, using Koss model UR 20 headphones, if participants were run in the lab, or on Vic Firth™ brand over-ear isolation headphones, to ensure presentation quality and isolation from external noise. Because stimuli were authentic recordings, they were presented without volume edits to preserve musicality. Participants adjusted volume to their comfort.

Trial stimuli were differentiated into three groups (n = 14), with each group representing a specific type of modulation: direct, pivot chord, and common tone. These three types of modulations were chosen to be maximally distinct. A fourth group of stimuli presented were lures that did not modulate (n = 7). The non-modulating stimuli were selected such that one excerpt from each composer was represented, except for Haydn, who was represented twice. Thus there were a total of 49 excerpts presented to participants, with a combined duration of 22m 59s.

In addition to being balanced across modulation types, test stimuli were also balanced across the three modulating conditions as to how many of each type changed mode during the modulation, as this was hypothesized to serve an obvious cue that a modulation has occurred. The excerpts were not balanced by starting mode, however, and more common tone excerpts started in the minor mode than any other. This is likely a negligible artifact, however, because the minor mode is not per se a cue that a modulation is going to happen.

Tempos for all stimuli were assessed to determine approximate tempo. In order to rule out any specific effects of tempo, we ran a simple post-hoc regression correlating *A*' and tempo, which was nonsignificant. Across all 14 stimuli, for each modulating condition, the average tempo was as follows: pivot chord, 112.64 beats per minute (bpm); direct, 115.00 bpm, and common tone, 114.50 bpm. Across the range of tempi included in the test stimuli, from a minimum of 43 bpm to a maximum of 236 bpm, the overall average was 114.05.

For each stimulus, we determined through aural and theoretical score analysis a critical time period in which the modulation or imitation lure occurred. Using that as a reference, we determined a time window in which a response would indicate an accurate reaction to, and therefore, perception of, the modulation. The window for each modulation began either on the modulation, as in the case of a direct modulation, or at the first indication of a modulation, which for the common tone excerpts was the beginning of the sustained note. The window for each modulation ended at the confirmation of the new key. For most excerpts this occured on a tonic chord of an authentic cadence in the new key. Most stimuli also had a designated window in which a response would count as a false alarm. These lure windows included such artifacts as secondary dominants[7], strongly emphasized minor chords, or non-harmonic tones, all of which resolved to the original tonic. None of these false alarm windows overlapped with the modulation windows and none could be considered true

[7] For example: V/V - V - I, read as “five of five going to five, going to one”. The “five of five” is the secondary dominant here - the dominant of the dominant.

modulations.

Key Distance

Traditionally in music, key distance is measured by the distance around the circle of fifths, illustrated in Figure **??**. This distance does a good job of measuring distance in terms of note differences, but doesn't accurately reflect the psychophysical or perceptual difference between keys, because, for example, a modulation from major to relative minor and major to its dominant would have the same distance, which does not account for the change in mode from major to minor, the added sharp note in the dominant key signature, or the effect of the intervallic distance between the tonics of two keys (Kleinsmith & Neill, 2018).

A more effective approach to mapping key distance is that of Krumhansl (1990), that served as the basis for Toiviainen and Krumhansl's 2003 study. Krumhansl used multi-dimensional scaling to map each of the keys, using their correlation profiles, (Krumhansl, 1990; Krumhansl & Kessler, 1982) onto a four-dimensional space. These were then mapped onto a torus, which was used by Toiviainen and Krumhansl (2003) as the spatial mapping for their concurrent probe-tone visualization. Using this model as a representation of a four-dimensional vector space, we were able to calculate the euclidean distance between each set of four coordinates. This method allowed us to determine a distance between each of the keys that accurately reflected the key correlation (i.e. the distance around the circle of fifths) and the psychophysical effects of mode change. Because each key is located at a specific point in four dimensional space, each of the calculated vectors between the keys is unique. However, there were some patterns that emerged. For example, the calculated distance between two keys separated by the interval a fifth was approximately 0.86. The smallest distance between any two keys was between a given tonic and its relative minor (e.g. C to a minor), approximately 0.65. The largest distance between any two keys was approximately 2, which was the distance between any two keys separated by a tritone (e.g. I - ♭V/ ♯IV; C to F♯/ G♭). The largest key distance for any excerpt in this experiment was 1.892, between c minor and A major. An ANOVA on the key distances of the stimuli indicated that there was not a significant difference in key distances among modulation types.

Procedure

Participants participated in the experiment at the MPaC lab at UT Dallas Main Campus, or another quiet, distraction-free environment as necessary. Following consent procedures, participants completed a questionnaire about the extent of their music training. Researchers then gave participants instructions on how to complete the task. Participants who were unfamiliar with the concept of a modulation received a brief introduction to the concept of tonality and modulation. Many participants who were not trained musicians were more comfortable with the term “key change”, and we used that connection to help those participants understand the overall concept. Once participants expressed a satisfactory understanding of the concepts, they were given instructions as to how to complete the task. Participants were then given a brief explanation of what they were to listen for, and that they should press the designated key on the keyboard when they hear the modulation within the stimulus. Participants were informed that they could respond as many times as they liked during any given excerpt, but that each excerpt only contained at most one modulation, and there were some excerpts that did not modulate. Participants were not informed in advance of what types of modulation to expect.[8] Stimuli were presented using Matlab version R2009B (Various, 2009) using code adapted from Raman and Dowling (2017). Responses in the modulation window were recorded as hits and responses in the false alarm window were recorded as false alarms. Responses outside of either window were evaluated as noise and not considered for the purposes of this analysis. Participants moved through the excerpts at their own pace, beginning each excerpt at their leisure following the completion of the

[8] A few highly-trained participants asked if they should consider the relative minor a key change. The response given was always “Respond when you think you’re in a new key”

previous one. Participants were allowed to take breaks as they felt necessary. Excerpts were presented in a different random order for each subject to mitigate any effects of ordering.

Design & Hypotheses

The planned analyses we considered for this experiment dealt with training, type of modulation, key distance, and response time. The first analysis we did was to find *A*' given hits and false alarms for all participants for all excerpts.[9] we then calculated average response time for each participant, and average time for each participant by modulation type, all of which were compared using planned Tukey tests with corrections for multiple comparisons.

Results

Training Level and Modulation type

A three by three mixed ANOVA comparing modulation type and training level with *A*' as the dependent variable indicated significant simple main effects of training, $F(2, 531) = 10.51$, $MSE = 0.02$, $p < .001$, and modulation type, $F(2, 531) = 59.48$, $MSE = 0.02$, $p < .001$, as well as a significant interaction, $F(4, 531) = 11.73$, $MSE = 0.02$, $p < .001$. A Tukey test indicated that there was a significant difference between the highly trained group and the nonmusician group, Cohen's d[10] = 0.41, 95% CI [0.21, 0.62], $p < .001$, as well as between the highly trained group and the moderately-trained group, $d = 0.26$, 95% CI [0.05, 0.47], $p < .001$. However, the difference between the moderately-trained group and the nonmusician group was not significant, $d = 0.15$, 95% CI [-0.05, 0.36], $p = .215$. These results are all depicted graphically in Figure 6.

The simple main effect of modulation type was significant, and a Tukey test indicated

[9] *A*' is an estimate of the unbiased proportion of correct responses where chance equals 0.50.

[10] Hereafter listed simply as *d*, not to be confused with the signal detection theory measure *d*'. In the current context, *d* represents the standardized mean difference between the means of the groups, calculated using the pooled standard deviation as a standardizer.

that the differences between each of the modulation types were all significant. Between direct modulation and pivot chord modulation was a difference of $d = 0.66$, 95% CI [0.45, 0.88], $p < .001$, between common tone modulation and pivot chord modulation was a difference of $d = 0.98$, 95% CI [0.76, 1.19], $p < .001$, between common tone modulation and direct modulation was a difference of $d = 0.31$, 95% CI [0.10, 0.52], $p = .002$. The results of the analysis by modulation type are all depicted in Figure 7, and within-group significant differences by modulation type are presented in Table 1.

Table 1

Significant Differences Among Training Levels and Modulation Types

	Cohen's d	lower limit	upper limit	p value
1 - CT vs. 1 - PC	0.69	0.48	0.90	.001
1 - CT vs. 1 - DM	0.69	0.48	0.91	$< .001$
2 - DM vs. 2 - PC	0.50	0.29	0.71	.041
2 - CT vs. 2 - PC	0.95	0.73	1.16	$< .001$
3 - DM vs. 3 - PC	1.49	1.26	1.72	$< .001$
3 - CT vs. 3 - PC	1.29	1.06	1.52	$< .001$

Note:

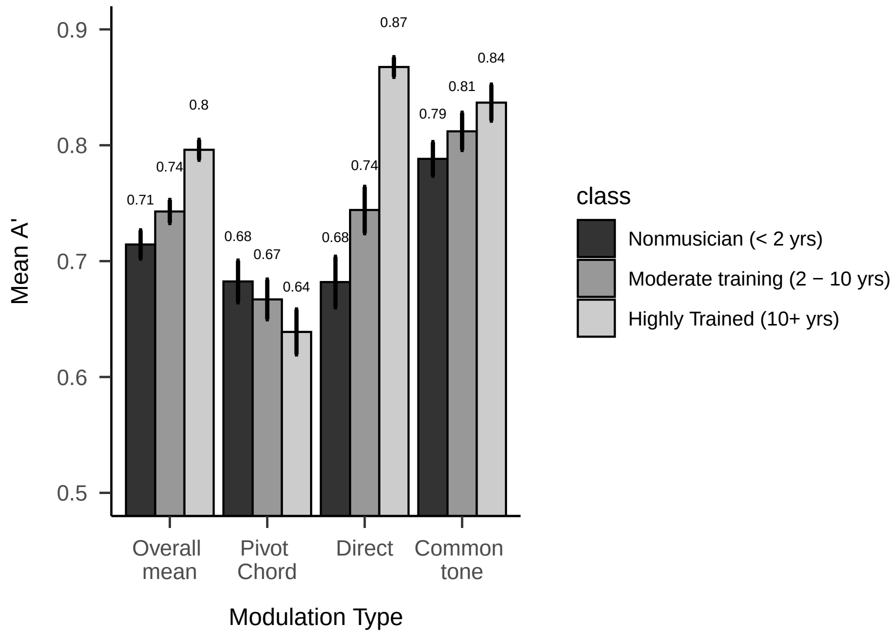
Groups: 1 = Untrained, 2 = Moderate training, 3 = Highly trained; Modulation types: CT = Common Tone, D = Direct, PC = Pivot Chord; Significance values after adjusting for multiple comparisons, lower limit and upper limit refer to the lower and upper limits of the 95% confidence interval for effect size.

Key Distance

A simple regression predicting A' from key distance was significant, $R^2 = .14$, 90% CI [0.02, 0.33], $F(1, 40) = 6.52$, $p = .015$, $R^2_{adj} = .12$, such that greater key distance indicated less accurate response, $b = -0.15$, 95% CI [-0.27, -0.03]. A multiple regression including both key distance and mode change was also found to be significant overall, $R^2 = .23$, 90% CI [0.03, 0.39], $F(3, 38) = 3.86$, $p = .017$. In this model, key distance was not found to be significant, $b = -0.04$, 95% CI [-0.22, 0.13], $t(38) = -0.49$, $p = .628$, but mode change was marginally significant, $b = 0.22$, 95% CI [-0.04, 0.49], $t(38) = 1.71$, $p = .096$, while the

Figure 6.

Relative accuracy of participants on the modulation perception task.



Note. Reported values of the A' statistic, showing both within and between - group effects. Error bars represent the standard error of the mean.

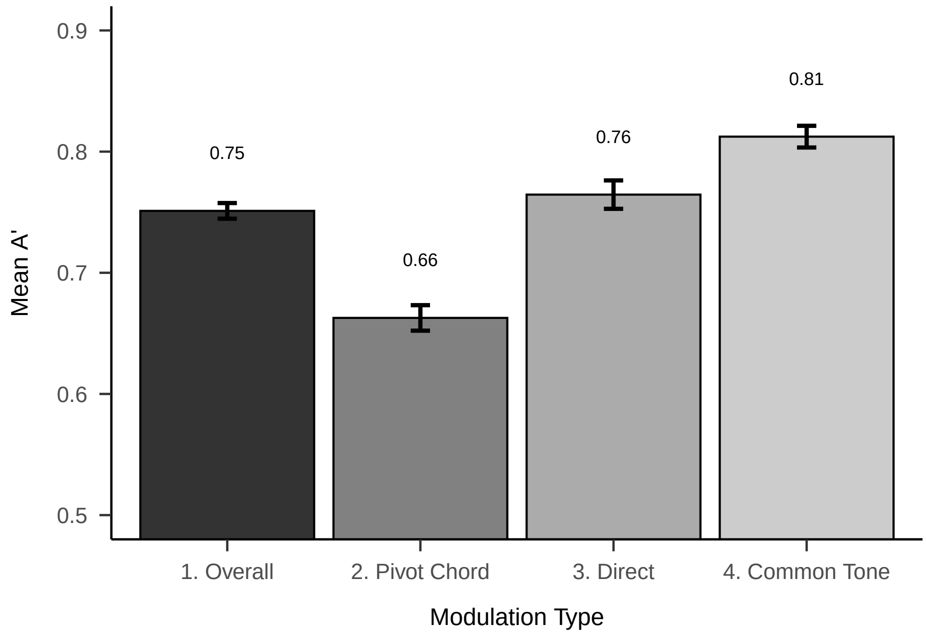
interaction between the two was found to be significant $b = -0.25$, 95% CI $[-0.49, 0.00]$, $t(38) = -2.03$, $p = .049$. These results are illustrated in Figure 8.

Reaction Time

Analysis of reaction time across training levels indicated a significant effect of training level, $F(2, 177) = 19.00$, $MSE = 0.21$, $p < .001$, $\hat{\eta}^2_G = .177$, and a Tukey test revealed significant differences between the highly trained musicians and non musicians group, $d = 0.65$, 95% CI $[0.44, 0.86]$, $p < .001$, and the highly trained group and the moderately-trained group, $d = 0.57$, 95% CI $[0.36, 0.78]$, $p < .001$, such that the highly trained musicians reacted more slowly to the modulations. However, the difference between the moderately-trained group and the nonmusican group was not significant, $d = 0.08$, 95% CI

Figure 7.

Relative accuracy of participants on the modulation perception task.



Note. Relative accuracy, represented by A', of participants on the modulation perception task, overall and modulation condition means. Error bars represent the standard error of the mean.

[-0.13, 0.28], $p = .789$. There was no difference in reaction time between modulation types, all $p > .05$. Finally, a regression predicting reaction time from key distance was not significant, $b = 0.20$, 95% CI $[-0.52, 0.91]$, $t(40) = 0.55$, $p = .585$. The results of reaction time by training are presented in Figure 9.

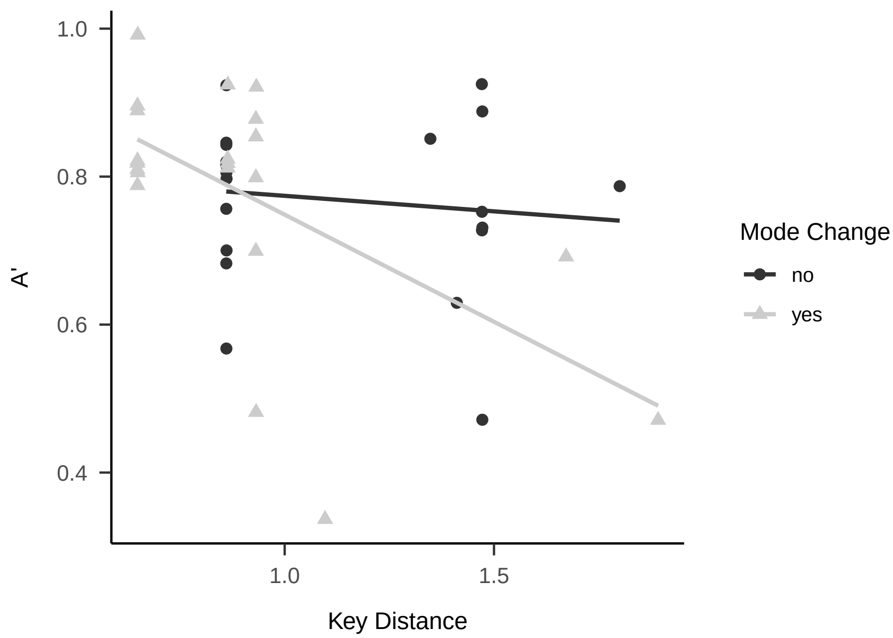
Discussion

In many respects, the results of the present study offer support for the existing hypotheses surrounding key distance and training, and in others it offers contradicting evidence. It also offers an interesting perspective on certain topical features that may help us understand what guides our understanding of tonic and tonic regions.

The significant results on response accuracy confirm a fairly common-sense idea that

Figure 8.

Relative accuracy of participants on the modulation perception task.



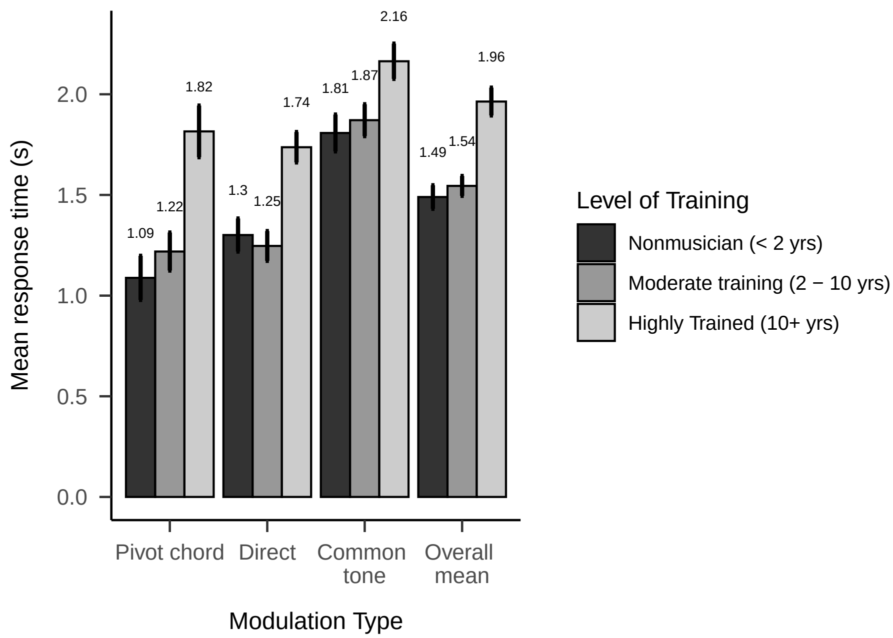
Note. Multiple regression using key distance and mode change as predictors.

training helps a listener more accurately discern if and when a modulation occurs. However, even the untrained listeners performed at above-chance levels on the task overall, suggesting that in general, listeners who are familiar with western music and its harmonic language are able to identify more often than not when a modulation occurs. This means that at some level, information regarding key area is encoded even in the minds of untrained listeners. Also, the fact that there was greater variablity in the accuracy of the trained participants than the untrained participants suggests that the quality of training is also important in terms of response accuracy. However, this variability may simply be due to individual differences in musical ability.

The majority of highly-trained listeners were in some way professional musicians, or training to be professional musicians; they are people for whom accuarate aural perception is

Figure 9.

Relative accuracy of participants on the modulation perception task.



Note. Graph representing the means the reaction times of the training groups. Error bars represent the standard error of the mean.

a professional necessity. It makes sense, then, that they would perform better on aural-skills tasks than untrained or moderately-trained listeners. It's also interesting that moderately-trained listeners, who have up to 10 years of formal training, are significantly less accurate than the highly trained listeners. This, of course, raises the question: are professional musicians better at aural skills tasks because they are professional musicians, or are they professional musicians because they are better at aural skills tasks?

The most interesting result in terms of modulation type is how poorly the highly trained listeners performed in the pivot-chord condition. It was by far the worst condition for the highly trained group and, although the differences were not significant, the highly trained group had the lowest mean score of the three groups. This phenomenon makes sense in light of what many of the participants said in their debriefing, namely that they weren't sure if many of the stimuli were true pivot chord modulations or if they were secondary dominants. Once they had made up their minds, it was too late; either the excerpt had ended or the window had passed (a fact which they would not have been aware of). This indecision is likely to have come from the instructions: participants were informed that there would be trials that did not modulate, and in their efforts to be accurate, they chose to not respond in some cases in which a response would have been correct. Specifically, musicians would be aware that one of the most prominent cues in the surface features of pivot chord modulations is also a prominent cue for a lure, in the form of a temporary tonicization, or the secondary dominant (V/V) described above. A temporary tonicization is therefore very similar to a pivot chord modulation in terms of both melodic features (altered or out of key notes functioning as leading tones) and harmonic features (altered chords functioning as secondary dominants). The primary difference between the two is whether or not the excerpt moves permanently to the new key, as in the case of a true modulation, or if it returns to the original starting key. As stated above, all of the *modulating* excerpts included in this study remain in the target key, but the fact that the excerpts were short may also have been a contributing factor. Although the new keys were confirmed by an authentic cadence in the new key, it's likely that this discrepancy is a behavioral difference, not a perceptual one, although that idea would require further testing. Pivot chord modulations led to the lowest A' score across all types of modulation. This makes sense when you consider the harmonic features of the pivot chord modulation. Incorporating the rare intervals theory described above (Butler, 1989), in any given key, the tritone between $\hat{7}$ and $\hat{4}$ is the most reliable predictor of a tonic region. Altering a note in the key seems to have the perceptual effect not of replacing that note, but expanding the set of notes in the perceptual window to include the chromatic alteration, at least until the intertia of perceptual experience in the piece erases the original note from the framework. None of these things happen in a purely melodic context, and incorporating the harmonic context gives us a clearer picture of this process. In general, the goal of composition in this idiom is smoothness, and the way to achieve that smoothness is by preceeding the altered note by a chord that assists in making the new note (and therefore the chord to which it belongs) appear “correct” (Figure 2). Thus, by expanding the tonic area to include the notes necessary to tonicize the new key, the composer enables the cadence in the new key to be effectively incorporated into the scheme without any perceptual jarring. However, those who are familiar with these cues will still recognize them for what they are. Dowling (1986) provides evidence that highly trained listeners encode pitches as scale steps when they’re listening to music, and this reflects that idea: the trained listeners were able to recognize the altered scale tones not only for their rank in the key set, but also for their function. Given the precision required for this task and subtlety of this compositional technique, it is very interesting that the untrained listeners performed as well as they did. Further analysis here would be warranted to investigate what other factors play into this result.

The greater spread between groups on the direct modulation more accurately exposes the effect of training. Whereas untrained listeners performed approximately as well on the direct modulations as the pivot chord modulations, both of the other groups performed significantly better on direct modulations than pivot chord modulations. Given the spread, it is likely that this effect is almost entirely dependent on training and has very little to do with surface features. People with music training of any kind are more familiar with the concepts of key and modulation, and are therefore consistently more accurate when responding, whereas those who are untrained are relying on whatever system seems to be tracking the tonic region.

For common tone modulations, all three training groups were clustered around $A' = .8$, which is by far the best performance overall. This result also supports the existing theories on pitch region perception and the surface features of the common tone modulations. These surface features align most closely with those of the probe-tone test paradigm, where the sustained or repeated note takes the place of the probe tone and serves as a reference pitch.

This result also contradicts our hypothesis that common tone modulations would be the second most accurately recognized modulation condition, after direct modulations. In creating the hypothesis, we was conflicted. we theorized that the effect could work one of two ways. Either the common tone would serve as a guide into the new key, helping listeners track pitch region and identify when the new key presented itself, or the common tone would obliterate the memory of the old key so that listeners would be unsure of what the old key was when they heard the new key. It seems that the first of those was correct. Also, since processing time has also been a factor in previous work on both melodies and modulations, (Raman & Dowling, 2017; Thompson & Cuddy, 1989) and one of the surface features of the common tone modulation is a long tone that allows the listener time to process the material, it makes sense that this would allow for the most accurate responses.

With regard to key distance, (Figure 8) it's important to look at the results considering the harmonic content of the stimuli. First of all, for stimuli that did not change mode, key distance had a negligible effect on overall accuracy. However, for stimuli that did change mode, there was a large effect of key distance, namely that stimuli that changed mode were far more likely to have more accurate responses the less distance they modulated and less likely to have accurate responses the greater the key distance, with A' clustering around .8 for excerpts that modulated to the relative minor, and falling to about chance for the greatest key distances. It is worth noting that there were only a few stimuli that modulated to distant keys, and better balancing of stimuli across key distances may shed more light on this effect. However, one interpretation of the results is that as modulation distance increases, mode change is more likely to act as a mask, obscuring the change in tonic, whereas stimuli that stay in the same mode have less musical material obscuring the modulation. These stimuli therefore maintain approximately the same level of accuracy regardless of key distance. Contrarily, stimuli that modulate shorter distances seem to get a boost from the mode change. This also makes sense in light of the current theories, that closer modulations are more easily recognized.

To incorporate some harmonic analysis into this discussion also helps illustrate this point: modulating to the relative minor is a fairly common modulation, and the harmonic distance between a given tonic and its relative minor is the smallest key distance possible. These stimuli that modulated to the relative minor were more likely to be recognized than even those that modulated to the dominant. At a glance, however, although the values predicted by the regression line for the non-modulating stimuli do not change very much, the graph shows a greater dispersion in A' values as key distance increases. These results of individual distances among stimuli are likely to come from factors not captured by this model, and would make for an interesting further investigation. This could examine effects of, for example, where the modulation occurs relative to a phrase boundary, or the complexity of the harmonic language in a given stimulus.

With regard to response time (Figure 9), our initial hypothesis that highly trained listeners would react faster to the modulations is not borne out by the data. Instead, trained listeners responded the slowest, across all types of modulation. We think the effect in this case may be similar to the effect seen with pivot chord modulations. Untrained listeners responded quickly to the modulations, an effect which comes perhaps from their reliance on their subconsious process or instincts. The untrained listeners thus suffered a higher false-alarm rate and a lower overall accuracy, which is reflected in their overall A' scores. Highly trained listeners, on the other hand, had recruited cognitive resources in accessing this information and used the processing time necessary to wait for confirmation that what they were hearing was, in fact, a modulation, as opposed to a temporary tonicization, and were thus more accurate. The paradox here is that the highly-trained participants were worse overall at accurately identifying pivot chord modulations, an effect which invites futher study. The data also seem to suggest that listeners with moderate training performed in a similar manner to the untrained listeners.

Exploratory Analyses

A post hoc exploratory Principal Components Analysis (PCA) performed on the stimulus data also showed some interesting results. The PCA included as variables: average beats per minute (bpm) and bpm range, reaction window start, end, and length, excerpt length, *tcrit*[11], time (as a part of the whole excerpt) before and after tcrit[12], key distance[13], date of composition, and *A*'. Permutation testing of the eigenvalues extracted by the PCA indicated that there were two significant dimensions extracted by the PCA. Dimension 1: $\lambda = 5.40$, $\tau = 41.55$, $p < .01$, Dimension 2: $\lambda = 2.42$, $\tau = 18.61$, $p < .01$. The results are depicted graphically in Figure 10.

Much of the tempo information defined the first dimension (horizontal), with average bpm in the positive direction strongly anti-correlated with tcrit, reaction window start and end, and excerpt length in the negative direction. This makes sense in that music with a higher average bpm, i.e. that is faster, will take less time to perform, assuming similar excerpt length in measures. The second dimension (vertical) was dominated by, in the positive direction, reaction window as a percentage of the total excerpt time, composition date, key distance, and time before the reaction window. These were all anti-correlated with *A*' in the negative direction on the second dimension. Bootstrap testing for consistency indicated that all of these variables loaded consistently on the dimensions with which they are associated, with three exceptions. Bpm range and time after tcrit did not load consistently on either dimension, and reaction window length loaded consistently on both dimensions. Because *A*' is on the second dimension and the tempo data is on the first, we interpret that to mean that *A*' is, in fact, uncorrelated/orthogonal with tempo across excerpts. However, the variables with which *A*' is anti-correlated support the results that we

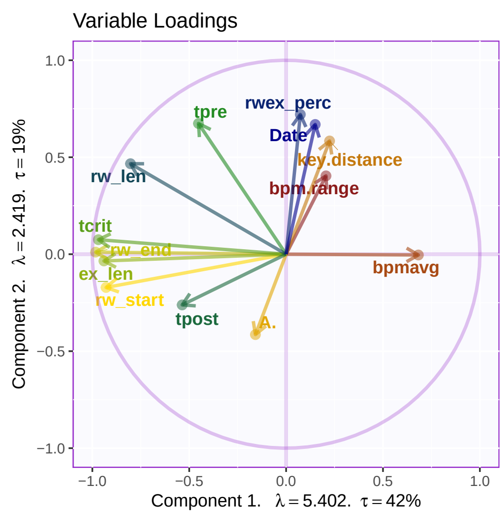
[11] tcrit is the precise moment of modulation, defined as the first appearance of the new tonic chord. For direct modulations, for example, this was equivalent to the beginning of the response window.

[12] All of the time variables listed here are measured in seconds.

[13] in arbitrary key distance units

Figure 10.

PCA variable loadings for exploratory analysis



Note. Plot represents the loadings of the variables included in the PCA. In this graph, the angles between the arrows indicate the strength and direction of their correlation. Arrows that are close to 0 and 180 degrees apart are strongly correlated and anti-correlated, respectively. Arrows that approach 90 degrees are uncorrelated, i.e. share no information; are orthogonal. The length of the arrow relative to the edge of the circle indicates how much of the variance of that variable was extracted and explained by the PCA. Lambda is the eigenvalue and tau is the percentage of variance extracted by that eigenvalue.

found in the regression above, that greater key distance correlates with lower *A*'. A novel revelation from this analysis is that Date, reaction window length, reaction window length as a percentage of overall excerpt length, and bpm range are all correlated with one another and are anti-correlated with *A*'. Interpreting this requires inference on our part. We know that music that was written later, in the Romantic or late Classical periods, tends to have more complex harmonic language. More complex harmonic language has two results: longer reaction windows, due to, for example, extended dominant area and longer cadential phrases, and more obscured tonic. The fact that tempo variability, represented by bpm range, is also in the mix here, has a few possible implications. One the one hand, tempo tends to vary more at phrase boundaries, which would lead us to expect that greater tempo variablity would be associated with an increase in response accuracy as participants respond to the ebb and flow of tempo at phrase boundaries. However, in terms of performance practice, tempo tends to be more constant for music from the classical era relative to that of the romantic era. The fact that A' prime is anti-correlated with bpm range in this model supports the effect of date, and by extension, era, in how participants perceived the modulations. This also shows that, at least in this model, A' is orthogonal to reaction window length, meaning that the length of the reaction window had no effect on the participant's response accuracy. This is interesting for two specific reasons: it suggests that harmonic language and complexity does in fact play a role in our ability to perceive modulations, and that a longer reaction window did not help participants to be more accurate.

It's curious here, though, that A' was significantly higher for excerpts featuring a common-tone modulation, which is a convention that only arose in the romantic era.

Future Directions

Future work in this vein should include excerpts that are more evenly balanced across key distance. Because this experiment focused specifically on surface features and modulation types, the effects of key distance may be overstated by the small sample size of large key distance modulations. Other topical effects that future research should attempt to rule out are phrase boundary effects and effects of harmonic language and complexity. Additionally, it would be interesting to look at cross cultural studies into other musical idioms and cultures, and to look at different age groups to analyze the effects of passive exposure to music over the lifetime. Most interesting, however, would be research into the cognitive lag question that arises from the reaction time results as well as the trained listeners' results on the pivot chord condition of the modulation type.

Conclusions

In summary, we found evidence for the following conclusions:

1. Listeners, across training levels, track tonic region independently of surface features.
2. Training helps, but only when that training is at or approaches a professional level.
3. The most helpful surface feature is a sustained pitch that both provides reference and time to allow for listener comprehension.
4. Trained listeners take longer to respond, but are more accurate, likely because they are analyzing the harmonic information in real time. This specific justification requires more in-depth study.
5. Prior evidence regarding key distance and modulation perception, specifically cognitive lag in processing greater key distance, is supported.
6. Highly trained listeners seem to be have conscious access to the information regarding pitch set content and the specific function of each pitch in the set.

References

Abdi, H. (2020). *PTCA4CATA: Partial triadic analysis for check all that apply (cata) data*. Retrieved from http://github.com/HerveAbdi/PTCA4CATA

Aust, F. (2019). *Citr: 'RStudio' add-in to insert markdown citations*. Retrieved from https://CRAN.R-project.org/package=citr

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*. Retrieved from https://github.com/crsh/papaja

Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(3), 501–515. https://doi.org/10.1037/0096-1523.6.3.501

Beaton, D., Fatt, C. R. C., & Abdi, H. (2014). An ExPosition of multivariate analysis with the singular value decomposition in R. *Computational Statistics & Data Analysis*, *72*(0), 176–189. https://doi.org/10.1016/j.csda.2013.11.006

Beaton, D., Fatt, C. R. C., & Abdi, H. (n.d.). An ExPosition of multivariate analysis with the singular value decomposition in R. *Computational Statistics & Data Analysis*, *72*(0), 176–189. https://doi.org/10.1016/j.csda.2013.11.006

Benjamin, T., Horvit, M., & Nelson, R. (2003). *Techniques and Materials of Music* (Sixth). Belmont, CA: Thompson.

Brown, H. (1988). The Interplay of Set Content and Temporal Context in a Functional Theory of Tonality Perception. *Music Perception: An Interdisciplinary Journal*, *5*(3), 219–249.

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: An Interdisciplinary Journal*, *6*(3), 219–241.

Dowling, W. J. (1978). Scale and Contour: Two Components of a Theory of Memory for Melodies. *Psychological Review*, *85*(4), 341–354.

https://doi.org/10.1037/0033-295X.85.4.341

Dowling, W. J. (1986). Context Effects on Melody Recognition: Scale-Step versus Interval Representations. *Music Perception*, *3*(3), 281–296.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (Third). Retrieved from https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Hebbali, A. (2020). *Olsrr: Tools for building ols regression models*. Retrieved from https://CRAN.R-project.org/package=olsrr

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

Kassambara, A. (2020). *Ggpubr: 'Ggplot2' based publication ready plots*. Retrieved from https://CRAN.R-project.org/package=ggpubr

Kausch, R. (2018). *Fre:Ac*. Retrieved from http://www.freac.org/

Kelley, K. (2020). *MBESS: The mbess r package*. Retrieved from https://CRAN.R-project.org/package=MBESS

Kleinsmith, A. L., & Neill, W. T. (2018). Recognition of transposed melodies: Effects of pitch distance and harmonic distance. *Psychonomic Bulletin and Review*, *25*(5), 1855–1860. https://doi.org/10.3758/s13423-017-1406-5

Krumhansl, C. L. (1983). Perceptual Structures for Tonal Music. *Music Perception: An Interdisciplinary Journal*, *1*(1), 33–57.

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Krumhansl, C. L., & Keil, F. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory & Cognition*, *10*(3), 243–251. Retrieved from https://link.springer.com/article/10.3758/BF03197636

Krumhansl, C. L., & Kessler, E. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368. https://doi.org/10.1037/0033-295X.89.4.334

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the Hierarchy of Tonal Functions Within a Diatonic Context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(3), 579–594.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music* (1st ed.). Cambridge, Massachusetts: MIT Press.

Müller, K., & Wickham, H. (2020). *Tibble: Simple data frames*. Retrieved from https://CRAN.R-project.org/package=tibble

Raman, R., & Dowling, W. J. (2016). Real-Time Probing of Modulations in South Indian Classical (Carnātic) Music by Indian and Western Musicians. *Music Perception*, *33*(3), 367–393. https://doi.org/10.1525/MP.2016.33.03.367

Raman, R., & Dowling, W. J. (2017). Perception of modulations in south indian classical (carnatic) music by student and teacher musicians: A cross-cultural study. *Music Perception*, *34*(4), 424–437.

R Core Team. (2020). *R: A language and environment for statistical computing*. Retrieved from https://www.R-project.org/

Schauberger, P., & Walker, A. (2020). *Openxlsx: Read, write and edit xlsx files*. Retrieved from https://CRAN.R-project.org/package=openxlsx

Sears, D. (2015). The Perception of Cadential Closure. In M. Neuwirth & P. Berge (Eds.), *What is a cadence* (pp. 253–286). Leuven, Belgium: Leuven University Press.

Thompson, W. F., & Cuddy, L. L. (1989). Sensitivity to Key Change in Chorale Sequences: A Comparison of Single Voices and Four- Voice Harmony. *Music Perception: An Interdisciplinary Journal*, *7*(2), 151–168.

Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception*, *32*(6), 741–766. https://doi.org/10.1068/p3312

Various. (2009). *Matlab R2009B*. The Mathworks, Inc.

Vuvan, D. T., Prince, J. B., & Schmuckler, M. A. (2011). Probing the Minor Tonal

Hierarchy. *Music Perception: An Interdisciplinary Journal*, *28*(5), 461–472. https://doi.org/10.1525/mp.2011.28.5.461

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12), 1–20. Retrieved from http://www.jstatsoft.org/v21/i12/

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, *40*(1), 1–29. Retrieved from http://www.jstatsoft.org/v40/i01/

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Retrieved from https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2020). *Dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Retrieved from https://yihui.org/knitr/

Appendix A

List of Recordings From Which Stimuli Were Excerpted

Beethoven, L. v. (1801). Op. 18, No. 1, Mvt. 4: Allegro [Recorded by Quartetto Italiano]. On Complete String Quartets [CD]. London: Decca. (1996)

Beethoven, L. v. (1801). Op. 18, No. 2, Mvt. 4: Allegro Molto quasi Presto [Recorded by Quartetto Italiano]. On Complete String Quartets [CD]. London: Decca. (1996)

Brahms, J. (1861). Op. 18, Mvt. 1: Allegro ma non troppo [Recorded by Amadeus String Quartet]. On Quintette; Sextette [CD]. W. Germany: Philips. (1968)

Brahms, J. (1865). Op. 36, Mvt. 4: Poco Allegro [Recorded by Amadeus String Quartet]. On Quintette; Sextette [CD]. W. Germany: Philips. (1968)

Brahms, J. (1873). Op. 51, No. 1, Mvt. 3: Allegro molto moderato e comodo [Recorded by Quartetto Italiano]. On The Complete String Quartets; The Complete Clarinet Sonatas [CD]. New York, NY: Philips. (1997)

Brahms, J. (1875). Op. 67, No. 3, Mvt. 1: Vivace [Recorded by Quartetto Italiano]. On The complete string quartets; The complete clarinet sonatas [CD]. New York, NY: Philips. (1997)

Brahms, J. (1890). Op. 111, No. 2, Mvt. 4: Vivace ma non troppo presto [Recorded by Boston Symphony Chamber Players]. On String Quintets [CD]. New York, NY: Elektra/Nonesuch. (1984)

Haydn, F. J. (1762). Op. 1, No. 1, Mvt. 1: Presto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1762). Op. 1, No. 1, Mvt. 3: Adagio [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1762). Op. 1, No. 1, Mvt. 5: Presto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1764). Op. 1, No. 2, Mvt. 1: Allegro Molto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1764). Op. 1, No. 2, Mvt. 2: Menuetto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1764). Op. 1, No. 3, Mvt. 1: Adagio [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1764). Op. 1, No. 3, Mvt. 2: Menuetto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1764). Op. 1, No. 3, Mvt. 4: Menuetto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1764). Op. 1, No. 4, Mvt. 3: Adagio [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1764). Op. 1, No. 4, Mvt. 4: Menuetto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 1, Nos. 1- 4. [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1764). Op. 1, No. 5, Mvt. 3: Allegro Molto [Recorded by Kodaly QUartet]. On Haydn: String Quartets Op. 1, Nos. 5 and 6 ; Op. 2, Nos. 1 and 2 [CD]. Hong Kong: Naxos. (1991)

Haydn, F. J. (1764). Op. 1, No. 6, Mvt. 5: Presto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Nos. 5-8, Op. 1, Nos. 0 and 6, and Op. 2, Nos. 1 and 2 [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1765). Op. 2, No 4, Mvt. 4: Menuetto: Allegretto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 42 and Op. 2, Nos 4 and 6 [CD]. Hong Kong: Naxos. (1993)

Haydn, F. J. (1765). Op. 2, No. 1, Mvt. 2: Menuetto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Nos. 5-8, Op. 1, Nos. 0 and 6, and Op. 2, Nos. 1 and 2 [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1765). Op. 2, No. 2, Mvt. 1: Allegro Molto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Nos. 5-8, Op. 1, Nos. 0 and 6, and Op. 2, Nos. 1 and 2 [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1765). Op. 2, No. 2, Mvt. 4: Menuetto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Nos. 5-8, Op. 1, Nos. 0 and 6, and Op. 2, Nos. 1 and 2 [CD]. Hong Kong: Naxos. (1992)

Haydn, F. J. (1765). Op. 2, No. 3, Mvt. 1: Allegro Molto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 2, Nos. 3 and 5 [CD]. Hong Kong: Naxos. (2003)

Haydn, F. J. (1765). Op. 2, No. 4, Mvt. 3: Adagio non troppo [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 42 and Op. 2, Nos 4 and 6 [CD]. Hong Kong: Naxos. (1993)

Haydn, F. J. (1766). Op. 2, No. 6, Mvt. 5: Presto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 42 and Op. 2, Nos 4 and 6 [CD]. Hong Kong: Naxos. (1993)

Haydn, F. J. (1769). Op. 9, No. 1, Mvt. 1: Moderato [Recorded by Kodaly Quartet]. On Haydn: String Quartets, Op. 9, Nos. 1, 3, and 4 [CD]. Hong Kong: Naxos. (1994)

Haydn, F. J. (1769). Op. 9, No. 1, Mvt. 3: Adagio [Recorded by Kodaly Quartet]. On Haydn: String Quartets, Op. 9, Nos. 1, 3, and 4 [CD]. Hong Kong: Naxos. (1994)

Haydn, F. J. (1769). Op. 9, No. 2, Mvt. 3: Adagio Cantabile [Recorded by Kodaly Quartet]. On Haydn: String Quartets, Op. 9, Nos. 1, 3, and 4 [CD]. Hong Kong: Naxos. (1994)

Haydn, F. J. (1769). Op. 9, No. 3, Mvt. 1: Allegro Moderato [Recorded by Kodaly Quartet]. On Haydn: String Quartets, Op. 9, Nos. 1, 3, and 4 [CD]. Hong Kong: Naxos. (1994)

Haydn, F. J. (1772). Op. 20, No. 2, Mvt. 1: Moderato [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 20, Nos. 1- 3, "Sun Quartets" [CD]. Hong Kong: Naxos. (1993)

Haydn, F. J. (1772). Op. 20, No. 6, Mvt 2: Adagio cantabile [Recorded by Kodaly Quartet]. On Haydn: String Quartets, Op. 20, Nos. 4-6, "Sun Quartets" [CD]. Hong Kong: Naxos. (1993)

Haydn, F. J. (1781). Op. 33, No. 2, Mvt. 1: Allegro Moderato, cantabile [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 33, Nos. 1, 2 and 5 [CD]. Hong Kong: Naxos. (1994)

Haydn, F. J. (1782). Op. 33, No. 5, Mvt. 2: Largo e cantabile [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 33, Nos. 1, 2 and 5. [CD]. Hong Kong: Naxos. (1994)

Haydn, F. J. (1790). Op. 64, No. 4, Mvt. 3: Adagio [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 64, Nos. 4-6. [CD]. Hong Kong: Naxos. (1993)

Haydn, F. J. (1790). Op. 64, No. 5, Mvt. 3: Minuet [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 64, Nos. 4-6. [CD]. Hong Kong: Naxos. (1993)

Haydn, F. J. (1790). Op. 76, No. 1, Mvt. 1: Allegro con Spirito [Recorded by Kodaly Quartet]. On String quartets, op. 76, nos. 1-3 [CD]. Hong Kong: Naxos. (1990)

Hofstetter, R. (1777). Op. 3, No. 1, Mvt. 1: Allegro Molto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 2, Nos. 3 and 5, Op. 3, Nos. 1 and 2 [CD]. Franklin, Tenn: Naxos. (2006)

Hofstetter, R. (1777). Op. 3, No. 1, Mvt. 3: Andantino Grazioso [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 2, Nos. 3 and 5, Op. 3, Nos. 1 and 2 (attrib. Hoffstetter) [CD]. Franklin, Tenn: Naxos. (2006)

Hofstetter, R. (1777). Op. 3, No. 1, Mvt. 4: Presto [Recorded by Kodaly Quartet]. On String Quartets Op. 2, Nos. 3 and 5, Op. 3, Nos. 1 and 2 (attrib. Hoffstetter) [CD]. Franklin, Tenn: Naxos. (2006)

Hofstetter, R. (1777). Op. 3, No. 3, Mvt. 3: Menutetto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 3, Nos. 3 - 6 [CD]. Hong Kong: Naxos. (2002)

Hofstetter, R. (1777). Op. 3, No. 5, Mvt. 1: Presto [Recorded by Kodaly Quartet]. On String Quartets: Op. 1, Nos. 5 and 6; Op. 2, Nos. 1 and 2 [CD]. Hong Kong: Naxos. (1991)

Hofstetter, R. (1777). Op. 3, No. 6, Mvt. 3: Menuetto [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 3, Nos. 3 - 6 [CD]. Hong Kong: Naxos. (2002)

Hofstetter, R. (1777). Op. 3, No. 6, Mvt. 4: Scherzando [Recorded by Kodaly Quartet]. On Haydn: String Quartets Op. 3, Nos. 3 - 6 [CD]. Hong Kong: Naxos. (2002)

Mozart, W. A. (1786). String Quartet No. 20 in D Major, KV 499 [Recorded by Quatuor Talich]. On Inte´grale des quatuor [CD]. France: Calliope. (1993)

Mozart, W. A. (1790). String Quartet No. 22 in Bb Major, KV 589 [Recorded by Quatuor Talich]. On Inte´grale des quatuors [CD]. France: Calliope. (1993)

Schubert, F (1812). String Quartet in Bb Major, D. 36, Mvt. 2: Andante [Recorded by Melos Quartett]. On The String Quartets [CD]. Hamburg: Deutsche Grammophon. (1999)

Schubert, F. (1814). String Quartet in Bb Major, D. 112, Mvt. 1: Allegro ma non troppo [Recorded by Wiener Konzerthausquartett]. On Les quatuors a‘ cordes [CD]. Paris: Universal Music, 1998. (1998)

Schubert, F. (1820). String Quartet Movement in C minor, D. 703 [Recorded by Melos Quartett]. On The String Quartets [CD]. Hamburg: Deutsche Grammophon. (1999)

Schubert, F. (1828). Quintet in C Major, Op. 163, D. 956, Mvt. 1: Allegro ma non troppo [Recorded by Bernard Greenhouse, Juilliard String Quartet]. On Quintet in C Major, Op. 163, D. 956 [CD]. New York, NY: CBS. (1988)

Schubert, F. (1828). Quintet in C Major, Op. 163, D. 956, Mvt. 2: Adagio [Recorded by Bernard Greenhouse, Juilliard String Quartet]. On Quintet in C Major, Op. 163, D. 956 [CD]. New York, NY: CBS. (1988)

Schubert, F. (1828). Quintet in C Major, Op. 163, D. 956, Mvt. 3: Scherzo [Recorded by Bernard Greenhouse, Juilliard String Quartet]. On Quintet in C Major, Op. 163, D. 956 [CD]. New York, NY: CBS. (1988)

Appendix B

List of Scores Consulted for Analysis

Beethoven, L. v. (1937). Op. 18, No. 1. New York: E. F. Kalmus Orchestra Scores. (Original work published 1801)

Beethoven, L. v. (1937). Op. 18, No. 2. New York: E. F. Kalmus Orchestra Scores. (Original work published 1801)

Brahms, J. (1927). Op. 18. Leipzig: Breitkopf & Härtel. (Original work published 1861)

Brahms, J. (1927). Op. 36. Leipzig: Breitkopf & Härtel. (Original work published 1865)

Brahms, J. (1927). Op. 51. Leipzig: Breitkopf & Härtel. (Original work published 1873)

Brahms, J. (1927). Op. 67. Leipzig: Breitkopf & Härtel. (Original work published 1875)

Brahms, J. (1927). Op. 111, No. 2. Leipzig: Breitkopf & Härtel. (Original work published 1890)

Haydn, F. J. (1845). Op. 1, No. 1. Berlin: Trautwein. (Original work published 1762)

Haydn, F. J. (1845). Op. 1, No. 2. Berlin: Trautwein. (Original work published 1764)

Haydn, F. J. (1845). Op. 1, No. 3. Berlin: Trautwein. (Original work published 1764)

Haydn, F. J. (1845). Op. 1, No. 4. Berlin: Trautwein. (Original work published 1764)

Haydn, F. J. (1845). Op. 1, No. 5. Berlin: Trautwein. (Original work published 1764)

Haydn, F. J. (1845). Op. 1, No. 6. Berlin: Trautwein. (Original work published 1764)

Haydn, F. J. (1845). Op. 2, No. 4. Berlin: Trautwein. (Original work published 1765)

Haydn, F. J. (1845). Op. 2, No. 1. Berlin: Trautwein. (Original work published 1765)

Haydn, F. J. (1845). Op. 2, No. 2. Berlin: Trautwein. (Original work published 1765)

Haydn, F. J. (1845). Op. 2, No. 3. Berlin: Trautwein. (Original work published 1765)

Haydn, F. J. (1845). Op. 2, No. 4. Berlin: Trautwein. (Original work published 1765)

Haydn, F. J. (1845). Op. 2, No. 6. Berlin: Trautwein. (Original work published 1766)

Haydn, F. J. (1845). Op. 9, No. 1. Berlin: Trautwein. (Original work published 1769)

Haydn, F. J. (1845). Op. 9, No. 1. Berlin: Trautwein. (Original work published 1769)

Haydn, F. J. (1930). Op. 9, No. 2. Leipzig: Ernst Eulenburg. (Original work published 1769)

Haydn, F. J. (1845). Op. 9, No. 3. Berlin: Trautwein. (Original work published 1769)

Haydn, F. J. (1930). Op. 20, No. 2. Leipzig: Ernst Eulenberg. (Original work published 1772)

Haydn, F. J. (1930). Op. 20, No. 6. Leipzig: Ernst Eulenburg. (Original work published 1772)

Haydn, F. J. (1930). Op. 33, No. 2. Leipzig: Ernst Eulenburg. (Original work published 1781)

Haydn, F. J. (1930). Op. 33, No. 5. Leipzig: Ernst Eulenburg. (Original work published 1782)

Haydn, F. J. (1968). Op. 64, No. 4. Moscow: State Publishers Music. (Original work published 1790)

Haydn, F. J. (1968). Op. 64, No. 5. Moscow: State Publishers Music. (Original work published 1790)

Haydn, F. J. (1968). Op. 76, No. 1. Moscow: State Publishers Music. (Original work published 1790)

Hofstetter, R. (1845). Op. 3, No. 1. Berlin: Trautwein. (Original work published 1777)

Hofstetter, R. (1845). Op. 3, No. 1. Berlin: Trautwein. (Original work published 1777)

Hofstetter, R. (1845). Op. 3, No. 1. Berlin: Trautwein. (Original work published 1777)

Hofstetter, R. (1845). Op. 3, No. 3. Berlin: Trautwein. (Original work published 1777)

Hofstetter, R. (1845). Op. 3, No. 5. Berlin: Trautwein. (Original work published 1777)

Hofstetter, R. (1845). Op. 3, No. 6. Berlin: Trautwein. (Original work published 1777)

Hofstetter, R. (1845). Op. 3, No. 6. Berlin: Trautwein. (Original work published 1777)

Mozart, W. A. (1882). String Quartet No. 20 in D Major, KV 499. Leipzig: Breitkopf and Härtel. (Original work published 1786)

Mozart, W. A. (1882). String Quartet No. 22 in Bb Major, KV 589. Leipzig: Breitkopf & Härtel. (Original work published 1790)

Schubert, F. (1973). String Quartet in Bb Major, D. 36. New York, NY: Dover Publications. (Original work published 1812)

Schubert, F. (1965). String Quartet in Bb Major, D. 112. New York, NY: Dover Publications. (Original work published 1814)

Schubert, F. (1965). String Quartet Movement in C minor, D. 703. New York, NY: Dover Publications. (Original work published 1820)

Schubert, F. (1965). Quintet in C Major, Op. 163, D. 956. New York, NY: Dover Publications. (Original work published 1828)