Testing Schenkerian theory: An experiment on the perception of key distances

Jason Yust,*[1]

* *School of Music, Boston University, USA*
[1]jyust@bu.edu

ABSTRACT

The lack of attention given to Schenkerian theory by empirical research in music is striking when compared to its status in music theory as a standard account of tonality. In this paper I advocate a different way of thinking of Schenkerian theory that can lead to empirically testable claims, and report on an experiment that shows how hypotheses derived from Schenker's theories explain features of listener's perception of key relationships.

To be relevant to empirical research, Schenker's theory must be treated as a collection of interrelated but independent theoretical claims rather than a comprehensive analytical method. These discrete theoretical claims can then lead to hypotheses that we can test through empirical methods. This makes it possible for Schenkerian theory improve our scientific understanding of how listeners understand tonal music. At the same time, it opens the possibility of challenging the usefulness of certain aspects of the theory.

This paper exemplifies the empirical project with an experiment on the perception of key distance. The results show that two features of Schenkerian theory predict how listeners rate stimuli in terms of key distance. The first is the Schenkerian principle of "composing out" a harmony, and the second is the theory of "voice-leading prolongations." In a regression analysis, both of these principles significantly improve upon a model of distance ratings based on change of scalar collection alone.

I. Schenkerian Theory in Music Perception Research

Schenkerian theory occupies a unique status between the disciplines of music theory and music psychology and cognition. Considered a standard account of tonal music by music theorists, it has only played a marginal role in the fields of music psychology and cognition. Only three abstracts in the over 30-year history of the journal *Music Perception* contain the words "Schenker" or "Schenkerian." Schenker's name never appears in abstracts from *Psychomusicology*, once in *Psychology of Music*, and twice in *Music Scientiæ*. By contrast it appears in 51 abstracts from the *Journal of Music Theory* and 43 in *Music Theory Spectrum*. (Searches conducted via Ebscohost.com, April 2012).

Music theorists have unintentionally discouraged empirical testing of Schenkerian theory by treating it as a comprehensive and integrated method of analysis, which is too complex to be tested directly. To be relevant to empirical research, Schenker's theory must be treated instead as a collection of interrelated but independent theoretical claims, each of which can lead to testable hypotheses. David Temperley (2011) argues a similar point from the perspective of corpus analysis.

Another significant reason that Schenker has been marginalized in music perception research is that his theories do not translate directly into experimental paradigms without a certain amount of sensitive interpretation. The project of testing Schenker therefore requires a high degree of interdisciplinary collaboration. Other more literal theories, such as the reductionism of Lerdahl and Jackendoff 1983, have tended to act as stand-ins for Schenker.

Experiments from a reductionist perspective have shown that listeners do hear melodies and harmonic progressions in terms of structural skeletons, and that reductions can serve to some extent as representations of structure (Dibben 1994, Serafine et al. 1989, Deutsch & Feroe 1981). They can also help to represent listeners' perceptions of musical tension in terms of long range melodic connections (Lerdahl & Krumhansl 2007).

One area where Schenkerian theory has been used more directly is in Steven Larson's theory of musical forces (Larson and McAdams 2004, Larson and VanHandel 2005), in which context it has proved essential in explaining listeners' expected continuations of incomplete melodies. Schenkerian analyses have also been shown to have some potential to predict similarity judgments on melodies (Martínez 2001).

II. Perception of Tonal Distance

Studies of key relationships have shown, not surprisingly, that listeners perceive key distance to some extent in the traditional music-theoretic sense of change of key signature, or distance on the circle of fifths. This traditional notion of distance can also be described in terms of change of scale: A larger change of key signature implies that the basic scales of two keys share fewer notes.

Studies involving the perception of phrases from Bach chorales have shown that listeners are sensitive to a change of key signature of up to two accidentals (Thompson & Cuddy 1989 & 1992). Yet listeners' ability to identify key signature change directly were poor in many instances, suggesting that other musical factors might also be at play. In studies by Krumhansl, Bharucha, and Castellano (1982) circle-of-fifths distance correlated with listeners' rating of chord similarity and confusions in the discrimination of transposed chord progressions.

Other studies have shown that while listener's are sensitive to key distances and the special status of the home key in musical passages, in certain paradigms they are insensitive to whether a musical passage begins and ends in the same key. (Marvin & Brinkman 1999, Cook 1987).

Temperley (2007) achieves a more sensitive measure of key distance by means of correlating key profiles, which can explain the likelihood of keys occurring in a piece of music from the common-practice era. For the purposes of perceptual ratings, such a measure will correlate closely with change of key signature.

III. An Experiment on Tonal Distance

The following experiment demonstrates the possibility of testing tenets of Schenkerian theory via a perceptual paradigm. The results show that listeners' ratings of key distances are strongly influenced by factors explained through Schenker's principles of composing out a harmony and voice-leading prolongations.

A. Design

Stimuli for the experiment consisted of three-voice homophonic chord progressions in a synthesized piano timbre. The first six chords established a key ("HK") ending in half cadence, the next six chords were in some contrasting scale ("CS"), and the last seven chords repeated the original progression ending on a PAC in the HK. Throughout the discussion here an in the data analysis, careted scale degree numbers and roman numerals should always be understood relative to the (major) HK.

Sixty-one undergraduate music majors at the University of Alabama participated in the study. They were divided into three groups who each heard a different series of 20 randomly generated chord progressions. They rated the perceived tonal distance in each progression on a scale from 0 to 10.

The stimuli were varied according to five parameters.

1. The CS varied from 0 to 6 accidentals flatward from the HK. Each progression was completely diatonic within the respective scales, following basic procedures of common practice harmony.
2. Two melodies were used in the HK context to outline either the space from $\hat{1}$ to $\hat{3}$ or $\hat{1}$ to $\hat{5}$.
3. The CS progression outlined a harmony whose root was a 5th below, a 3rd below, or 2nd above the HK tonic.
4. The CS melody, which was always a stepwise descent through a sixth, started from either the root or third of the outlined chord (and hence would end on the third or fifth, respectively).
5. The HK was varied between E, F, F#, and G major.

See the sample stimuli in Fig. 1 (a), (c), and (e).

B. Hypotheses Derived from Schenkerian Theory

Three factors were expected to predict distance ratings:

First, a larger change of scale was expected to lead to higher ratings, in accordance with the traditional theory of keys and previous findings.

Second, listeners were expected to generalize the progressions in terms of "composed-out" harmonies, which would be the HK tonic in the framing progressions, and the outlined chord in the CS progression. Therefore, the more logical the harmonic relationship between chords (by fifth first, and by third next, and by step last), the lower the expected rating. Notice that the outlined harmony in the CS was *not* necessarily the tonic of a traditional major or minor key, so harmonic relationship and scalar relationship were essentially independent (the only caveat being that the scale could not be chosen in such a way that the outlined chord was diminished or the chord with a root a fourth above it, which always featured prominently in the progression, was diminished. Or, to put it differently, the CS might be thought of as one of five modes, Ionian, Dorian, Phrygian, Mixolydian, or Aeolian, relative to the outlined chord.)

This second hypothesis, though characteristically Schenkerian, is not exclusively Schenkerian. Given the construction of the stimuli, an effect of this nature might be predicted by other assumptions. For instance, one might hypothesize that listeners respond to the prominence of the outlined chord in the phrasing of the stimuli and rate its relationship to an established context. Further research will be needed to fully disambiguate the most characteristically Schenkerian explanation for this prediction.

Figure 1. Three sample stimuli and interpretations. In (a), a 4-accidental shift. Under the Schenkerian prediction, (b), this is heard harmonically as a third relationship, and the verticalized intervals outlined by the melody created an upper neighbor figure. In (c)–(d), another four-accidental shift, but here to a fifth-related chord. The verticalized intervals suggest a passing motion that could be interpreted as a "motion to an inner voice," a less obvious melodic relationship than the upper neighbor. In (e)–(f), a three-accidental shift to a fifth-related chord. The verticalized melodic intervals suggest no obvious melodic relationship between the outlined harmonies.

The third hypothesis is the one of greatest present interest: that the Schenkerian principle of voice-leading prolongations will help explain distance ratings. According to this principle, we would expect listeners to summarize the three parts of the progression not only in terms of the scale or outlined chord, but also in terms of the specific interval that is outlined in the melody. These outlined intervals imply a voice leading between the composed-out harmonies in the three parts of the progression. This voice leading is simple when the first note of the CS progression is an upper neighbor to the highest note in the HK progression, less simple when the first note of the CS progression is a lower neighbor (Schenkerian theory would interpret this as "motion to an inner voice," which is not as characteristic of a melodic derivation of flat-side keys), and extravagant when the first note of the CS progression is a fourth above the highest note of the HK progression. Fig. 1 (b), (d), and (f) illustrates the three possibilities, for three sample stimuli.

C. Results

In the following discussion I will refer to the melodic conditions as "leap," "LN," and "UN," and the harmonic conditions as "I–IV," "I–VI," and "I–II" (or "IV," "VI," etc.).

The design was not balanced and hence was analyzed using a multiple regression model. Two regressions were run: a simple model with main effects only and a full model with interactions. The melodic and harmonic conditions were dummy coded with IV and UN (the most frequent and lowest rated) conditions as baseline. The regression treated key distance as a numerical variable and included subjects factors. In the simple model coefficients main effects in key, leap, and I–II were highly significant, while the effect of I–VI was marginal (p=.03). In the larger model all main effects were significant except I–II and there were significant interactions between key distance, the LN condition, and the chord conditions, including a three-way interaction. The criterion for significance was p < .001 in all cases. Another regression was run with only the significant factors from the large model. Table 1 shows the coefficients from each regression.

The design was imbalanced not only because stimuli parameters were selected by a randomized function, but also because of logical constraints on the experimental conditions. In the LN condition the CS melody had to start from $\hat{2}$ or $\hat{4}$, so it could not outline a VI. Similarly, the II and leap conditions could not occur together. Therefore, in the full model, interactions involving VI and LN or UN were not present simply because the relevant conditions did not exist.

Table 1. Regression coefficients for a simple model (main effects only) and a full model with interactions. All effects shown are significant at p < .001. "Key" is a numerical variable (its coefficient represents the change of rating associated with a one-accidental change in key signature). Other conditions are dummy coded with "UN" and "I-IV" as baselines.

	Simple model (Adjusted R^2 = .208)		Full Model (Adjusted R^2 = .264)	
Factor	**Coefficient**	**Std. Err.**	**Coefficient**	**Std. Err.**
Intercept	1.69	0.52	0.98	0.51
Key	0.48	0.04	0.79	0.06
LN			2.76	0.43
leap	1.33	0.15	1.15	0.14
I–VI	(0.32)	(0.15)	1.84	0.28
I–II	1.51	0.27		
key*LN			–0.98	0.14
key*VI			–0.50	0.08
key*II			1.79	0.24
key*LN*II			–1.38	0.27

Fig. 2 shows the effect of key distance overall. Figs. 3–4 show the overall responses for the harmonic and melodic conditions, respectively. Fig. 5 separates out the key distance effect in the LN condition, for interpretation of the key*LN interaction. Fig. 6 shows the effect of the melodic conditions after factoring out key distance and the key*LN interaction, by plotting the residuals of a regression against the three melodic conditions. Fig. 7 separates out the key distance effect in the three harmonic conditions for analysis of the interactions between key and harmony.

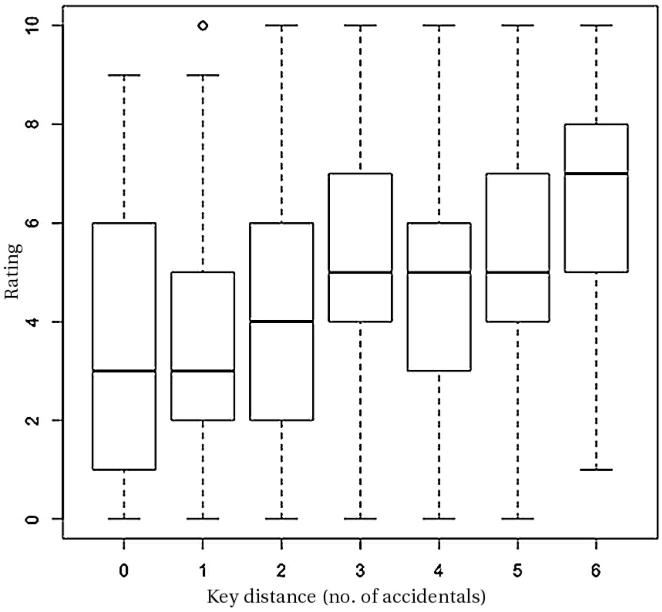

Figure 2: Rating by key distance.

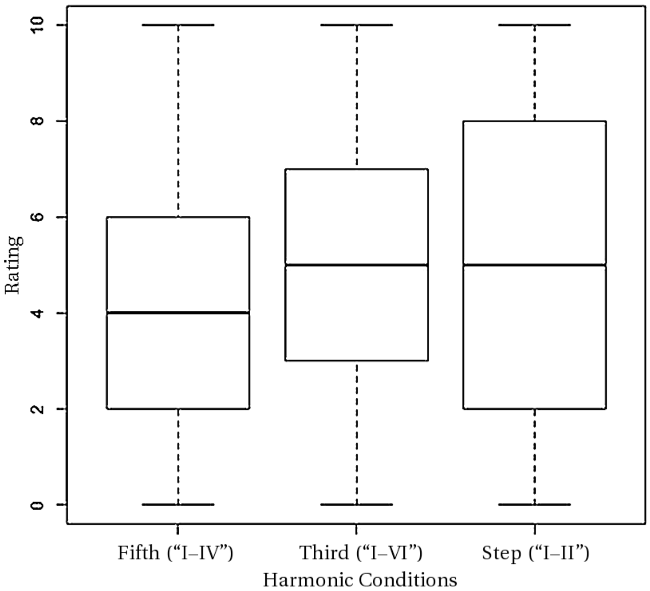

Figure 3: Rating by harmonic conditions.

D. Analysis of Results

The simple regression broadly confirms the hypotheses. First, key signature changes had a reliable effect on distance ratings, corresponding on average to a rating one point higher for each two accidentals in the key signature. The chord conditions also had the predicted effect, with a lower rating for more closely-related chords. The magnitude of this effect was considerable when compared with the key signature, the closer harmonic relationship of the I–IV condition lowering distance judgments in Fig. 3 to a similar extent that a two-accidental change in key signature does in Fig. 2. The same is true of the melodic conditions, where the disconnected melodies increased ratings by about the same

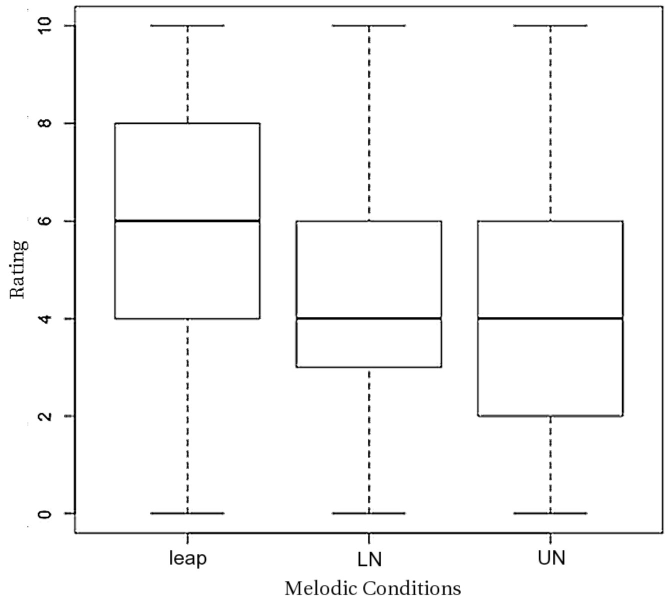

Figure 4: Rating by melodic conditions.

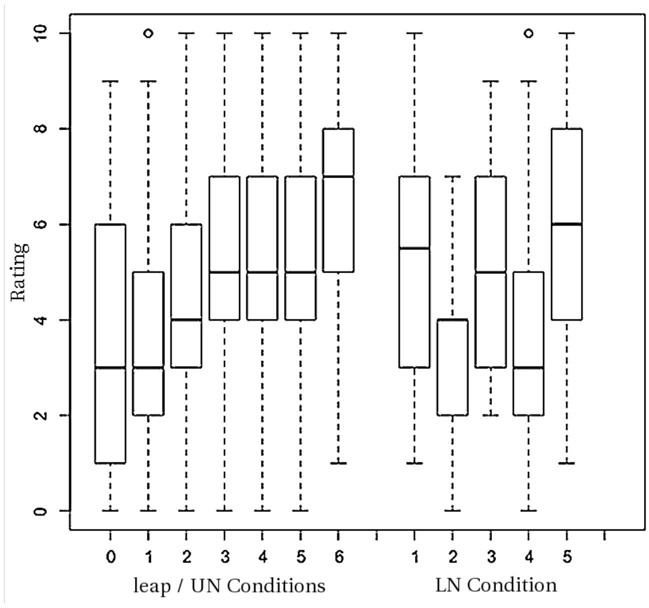

Figure 5: Rating by key distance in two melodic conditions.

rate as three accidentals in the key signature. This is remarkable considering that keys just two accidentals apart are considered distant and are avoided altogether in some musical styles (early 18th century music, for instance).

The full model shows that the effects of the experimental conditions were not simple independent influences on subjects ratings. The melodic and harmonic conditions were significant partly because of such simple main effects and partly by changing the way that key distance influenced ratings.

First, the primary effect of LN was to essentially cancel the effect of key distance on ratings. The sum of the interaction coefficient and the key coefficient is close to zero, and the high coefficient on LN simply means that the effect of key distance was replaced by a constant rating somewhere in the middle of the response scale. This is illustrated in Fig. 5,

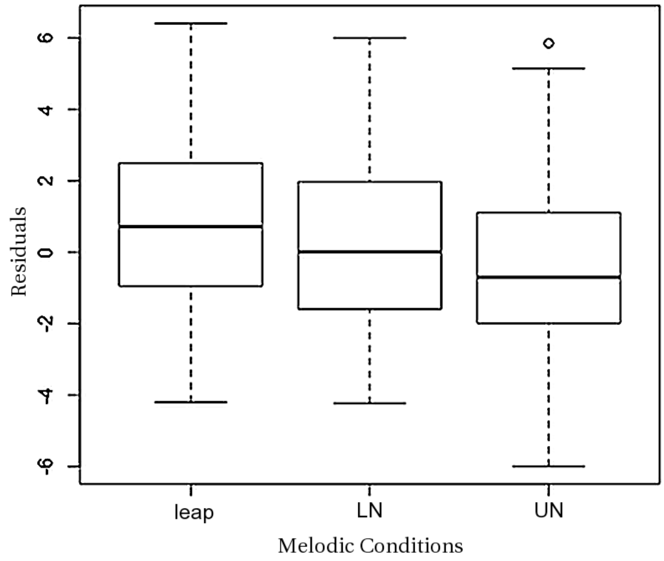

Figure 6. Residuals of a regression on subjects, key distance, and the interaction of key distance with the LN condition in the three melodic conditions.

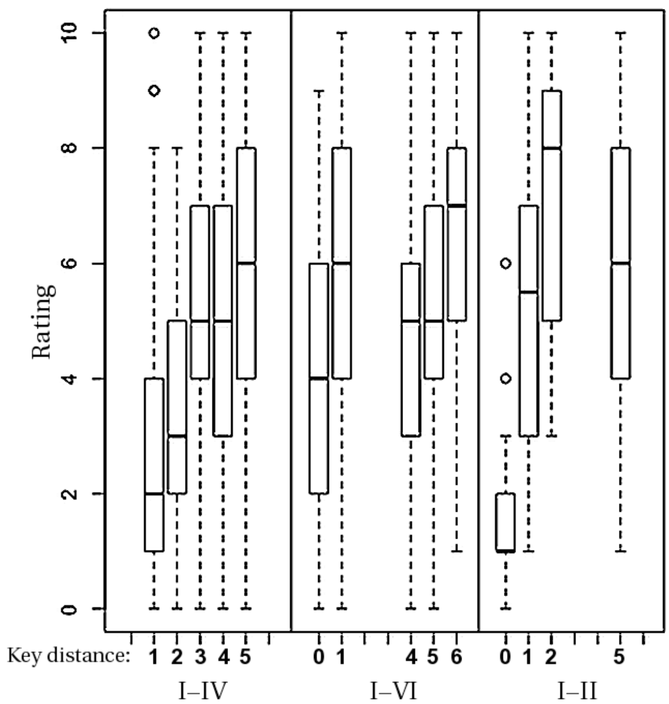

Figure 7: Rating by key distance in three harmonic conditions.

which separates out the key distance effect in the LN condition.

The nullifying effect of the LN condition makes sense in that the CS melody would never start from $\hat{6}$ in the LN condition, and would instead start from a note, $\hat{4}$ or $\hat{2}$ that was unlikely to be altered. If the listener's attention were drawn to the large-scale melodic progression as hypothesized, the changed accidentals would be "hidden" in an inner voice or a foreground passing line. The exception to this would be stimuli in the LN and I–II conditions at a key distance of 5 (so the CS melody outlined a ♭II chord starting from ♭$\hat{2}$). High ratings for this specific stimulus would be accounted for in the regression by coefficients associated with the I–II condition.

The overall effect of the melodic conditions after controlling for the key*LN interaction is illustrated in Fig. 6, which controls for key and harmonic factors as well as the interaction by running a regression on just these factors and plotting the residuals of this regression against the melodic conditions. The overall effects of these factors fit the hypothesis derived from Schenkerian theory, with the UN conditions leading to relatively low ratings, the leap condition to relatively high ratings, and the LN condition to average ratings.

The interactions of harmonic factors with key distance can be understood by examining the data in Fig. 7. When the CS progression outlined a IV, key distance has an approximately linear effect on responses. However, in the I–VI and I–II conditions the effect seems to be complicated by the contradictory effect of mode: For key distances 4 accidentals or higher, the CS progressions would outline a major triad (♭VI or ♭II), and there may have been some tendency to give lower ratings where the CS progression outlined a major triad instead of a minor triad in spite of key distance. In the I–VI condition the interaction factor therefore simply cancels out most of the effect of key distance. In the I–II condition, on the other hand, the regression model was able to separate the key distance = 5 stimuli from key distance = 0, 1, or 2 on the basis of the LN factor, because large key distances in the I–II condition only occurred (coincidentally) in the LN condition. Therefore the model was able to include large key distance effect for the I–II condition and cancel it out for key distance = 5 through the three-way interaction term.

IV. Conclusion

The results of the key distance experiment generally support the hypotheses based on Schenkerian theory, particularly the finding that a disconnected background melody increases key distance ratings as much as a 2–3 accidental change in the key signature. The experiment also found a difference between flat-side keys anchored by a background upper neighbor motion and a background melody suggesting "motion to an inner voice," with the latter leading to lower ratings overall, but also making changes of scale less perceptible. The weakened effect of key distance for "motion to an inner voice" stimuli supports the idea of hearing in terms of a background melody, since the upper neighbor and leapy background melodies would include the altered notes of the contrasting key whereas the "motion to an inner voice" background melodies usually would not.

Finally, the result showed in a number of ways that listeners hear a progression in terms of structural chords, making it plausible that relationships between structural chords are more important than changes of scale alone in listener's evaluation of key distance. This was in evidence in the lower distance ratings when structural chords were closely related, having an effect comparable to a two-accidental change of key signature.

Finally, and most importantly, the experiment shows that Schenkerian theory can be used to construct research paradigms that advance our understanding of music perception. The perception of key distance is one promising avenue for further such research, and the Schenkerian theories of composing-out and voice-leading prolongation, which informed the design of the present experiment, can be used similarly to construct paradigms that test other aspects of the listener's experience of tonal music.

ACKNOWLEDGMENT

Thanks to Thomas Robinson for helping provide subjects and run the experiment.

REFERENCES

Cook, N. (1987). The perception of large-scale closure. *Music Perception* 5(2), 197–206.

Deutsch, D. & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review* 88(6), 503–522.

Dibben, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *Music Perception* 12(1), 1–25.

Krumhansl, C., Bharucha, J. J. & Castellano, M. (1982). Key distance effects on perceived harmonic structure in music. *Perception & Psychophysics* 32(2), 96–108.

Larson, S. & McAdams, S. (2004). Musical forces and melodic expectations: Comparing computer models and experimental results. *Music Perception* 21(4), 457–498.

Larson, S. & VanHandel, L. (2005). Measuring musical forces. *Music Perception* 23(2), 119–136.

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

Lerdahl, F. & Krumhansl, C. (2007). Modeling tonal tension. *Music Perception* 24(4): 329–66.

Martínez, I. C. (2001). Contextual factors in the perceptual similarity of melodies. *Mikropolyphonie* 7.

Marvin, E. & Brinkman, A. (1999). The effect of modulation and formal manipulation on perception of tonic closure by expert listeners. *Music Perception* 16(4), 389–407.

Serafine, M. L., Glassman, N., & Overbeeke, C. (1989). The cognitive reality of hierarchic structure in music. *Music Perception* 6(4), 397–430.

Temperley, D. (2007). Key relations. In *Music and Probability* (Cambridge, MA: MIT Press), 99–108.

——— (2011). Composition, perception, and Schenkerian theory. *Music Theory Spectrum* 33(2), 146–168.

Thompson, W. F. & Cuddy, L. (1989). Sensitivity to key change in chorale sequences: A comparison of single voices and four-voice harmony. *Music Perception* 7(2), 151–168.

Thompson, W. F. & Cuddy, L. (1992). Perceived key distance in four-voice harmony and single voices. *Music Perception* 9(4), 427–438.