Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Closure and Expectation: Listener Segmentation of Mozart Minuets

Crystal A. Peebles

School of Music, Northern Arizona University, United States
Crystal.Peebles@nau.edu

ABSTRACT

This study investigates the theoretical claim that the perception of closure stems from the ability to predict the completion of a schematic unit, resulting in a transient increase in prediction error for the subsequent event. In this study, participants were asked to predict the moment of completion of mid-level formal units while listening to three complete minuet movements by Mozart (K. 156, K. 168, and K. 173). Following this prediction task, participants then rated the degree of finality of ending gestures from these same movements. Generally, endings punctuated by strong cadential arrival were best predicted and received higher ratings, suggesting that learned harmonic and melodic ending gestures contribute to the segmentation of musical experience. These results were accentuated for participants with formal musical training, further supporting this conclusion.

I. BACKGROUND

Listeners and musicians have intense aesthetic opinions regarding the degree to which a particular musical ending sounds satisfying, and “closure” is the oft-used term to describe the listener’s feeling of finality or completeness. The concept of musical closure is ubiquitous in the musicological discourse: it can steer analytical choices, propel a narrative trajectory, or provide scaffolding for a theoretical construct. Despite these different uses of “closure,” similarities in closural metaphors speak to our shared experience of finality. For instance, closure is often described as the experience of finality marked by the conclusion of a goal-directed process that segments musical experience into discrete units (Meyer, 1956; Meyer 1973; Agawu, 1987; Narmour, 1990; Snyder, 2000; Huron 2006; Anson-Cartwright, 2007). This understanding of closure (closure marks the achievement of a musical goal and segments a listener’s musical experience) highlights the importance and interrelatedness of expectation and segmentation in the perception of closure.

The connection between expectation and segmentation has been explored in the segmentation of every-day situations into meaningful events. According to the Event Segmentation Theory (EST), “perceptual systems spontaneously segment activity into events as a side effect of trying to anticipate upcoming information” (Kurby & Zacks, 2007, 72). In this theory, both unanticipated changes in sensory input and goal completion, as determined by pre-existing knowledge structures, create meaningful events out of a continuous stream of activity (Zacks, 2004). While this model in its entirety has not been explicitly tested with music, various claims of this model have. EST posits that segmentation is automatic and does not require conscious attention (Knösche, et al. 2005) and that segmentation occurs on multiple times scales simultaneously, creating a hierarchical grouping structure (Krumhansl, 1996).

More recently, Pearce, Müllensiefen, and Wiggins (2010) illustrate that expectation, based on probabilistic learning, informs the segmentation of musical melodies. They hypothesize that, similar to the mechanisms of segmentation described by EST, “boundaries are perceived before events for which the unexpectedness of the outcome and the uncertainty of the prediction are high” (1375). The first principle of segmentation, “the unexpectedness of the outcome,” refers to unanticipated discontinuities in the sound. Their results indicate that a computer model, using only probabilistic learning, can segment a melody according to their first principle in a way that mirrors the results from expert listeners.

While discontinuities in sound certainly do segment music from a phenomenological perspective (Lerdahl & Jackendoff, 1983; Deliège, 1987), many analysts would hesitate to associate this retrospective marking of a boundary with “closure”. For instance, Hopkins (1990), following in Meyer’s footsteps (1956, 1973), states, “for closure to occur, it is necessary . . . for a discernible process or pattern in one or more musical parameters to imply a particular point of conclusion” (4). From the perspective of probabilistic learning, the high transitional probabilities between successive events in a cadential gesture would allow a listener to anticipate the moment of conclusion, resulting in a boundary with a different phenomenological experience than one based on sonic discontinuities alone (Huron, 2006).

Specifically consider the harmonic paradigm for a perfect authentic cadence in the Classical style. As the cadence approaches, a hypothetical listener familiar with the Western tradition would be able to predict with certainty the tonic arrival due to this chord’s high transitional probability following a dominant chord. The transitional probability of this tonic chord going to any particular harmony is low compared with the probability of the dominant chord moving to tonic, resulting in a transient increase in prediction error. This increase of uncertainty following a cadence would also contribute to the sense of finality (Huron, 2006). It is this uncertainty, according to EST and the second principle of segmentation by Pearce, Müllensiefen, and Wiggins (“boundaries are perceived before events for which . . . the uncertainty of the prediction [is] high”) that prompts the creation of a perceptual boundary and the perception of closure.

This study investigates the theoretical claim that anticipated endings followed by a rise in uncertainty for the subsequent musical event correspond with the formation of a perceptual boundary and the feeling of finality. While this study does not directly measure the hypothesized transient increase of prediction error, it does measure the ability of participants to predict phrase endings in three complete minuet movements by W.A. Mozart (K. 156, K. 168, and K. 173). This relies on the assumption that the better a listener can anticipate a particular ending, the steeper the drop in the listener’s ability to predict the subsequent event. A listener’s rating of the perceived finality of a particular ending should then correlate with the predictability of that ending.

II. METHOD

A. Participants

Participants in this study were divided into three subject groups determined by their levels of formal musical training. Data from 24 undergraduate non-musicians (who received psychology credit for participating in this study), 27 undergraduate musicians (who received extra credit in their freshman Music Theory class), and 23 graduate student musicians (who received a $10 gift card) were included in this study.

B. Stimuli

Stimuli included minuet and trio movements from three string quartets by W.A. Mozart: String Quartet No. 3 in G major (K. 156), third movement; String Quartet No. 8 in F major (K. 168), third movement; and String Quartet No. 13 in D minor (K. 173), third movement. These three particular movements were selected for this study because each exhibits mostly conventional phrase rhythm with a few phrase expansions that disrupt the usual four-measure groupings. This variability in the length of phrases would force the listener not to rely exclusively on metrical cycles to anticipate endings. For example, K. 156 features a prominent deceptive cadence in m. 58, initiating a four-measure phrase expansion, and K. 168 features a dominant harmonic arrival in m. 15, which is then extended an additional measure.

Taken together, these movements include all three significant cadence types: perfect authentic cadence (PAC), imperfect authentic cadence (IAC), and half cadence (HC). In this study, a cadence is defined as a harmonic and melodic paradigm that concludes a mid-level formal unit (Caplin, 2004). Harmonically, both a PAC and an IAC contain a root-position dominant harmony that resolves to a root-position tonic harmony, but a PAC must end with scale-degree 1 in the melody, while an IAC can conclude with either scale-degrees 3 or 5. A root-position dominant harmony marks the arrival of a HC. Most of the cadences in these three movements do not involve a melodic suspension, meaning the harmonic and melodic arrivals coincide.

The clips used in the rating task represent subphrase, phrase, and section endings. For the purposes of this study, a subphrase is defined as a formal unit shorter than a phrase that has a beginning and end, which is not necessarily marked by a cadence, while a phrase is a longer formal unit that does conclude with a cadence. A section of music is determined by the macro-form of the movement (in the case of all three movements here, a composite ternary). The clips varied in length from two to six measures, where each clip began with the onset of a formal unit (subphrase or phrase) and concluded with the release of the last sound of that formal unit.

For the prediction task, all three movements, as performed by the Amadeus Quartet, were combined into a single audio file with fifteen seconds of silence between successive movements (the order of the movements were counterbalanced between participants). For the rating task, the clips were created by splicing the original audio file using the digital editing software Audacity.

C. Procedure

After participants gave informed consent, they were read the instructions for the prediction part of the study: "In the first part of the study you will listen to several pieces. While listening, try to predict the moment at which a musical phrase ends. Your goal is to press the spacebar at the exact moment the phrase is completed. The composer could surprise you, so it's okay if you press prematurely. Just keep on listening and try to anticipate the next ending." For participants unsure of the definition of a musical phrase, a linguistic analogy was used. It was explained that in instrumental music, notes group together to form sentence-like units, and their job is similar to predicting the beginning of the last word of a sentence. To ensure understanding of the prediction task, participants performed the task on two practice excerpts: the first reprise from Mozart's String Quintet No. 4 in G minor, third movement (K. 516) and Mozart's Sonata for Piano and Violin in B-flat major, third movement, mm. 1–16 (K. 454). These two excerpts were chosen because they illustrated the nature of the task, training the participants to predict ending actively rather than react retrospectively to a phrase ending.

At a personal computer, participants listened to the musical examples through headphones played through the software program Audacity. As the music progressed, participants indicated their predictions by pushing a key on the keyboard and Audacity recorded the time of the press. Following the experiment, the response times were exported as a text file for data analysis.

When participants had finished the prediction task, they began the rating task. The directions read: "In this part of the study you will use a seven-point scale to rate how well the short musical clip would end a musical idea. In music, some endings sound more conclusive than others. If the musical clip does not sound at all like an ending, then press 1. Use higher numbers to indicate stronger endings, with 7 representing the strongest possible ending. Since this is a matter of opinion, don't worry that there is a right or wrong answer. Feel free to use the entire range of the scale." Using the same computer, participants listened to each clip once and then selected their rating. Both the prediction and rating tasks were completed without reference to a score, and the entire experiment took between 30 and 45 minutes. Following the completion of both parts, participants filled out a questionnaire documenting their familiarity with the compositions and their musical training.

III. RESULTS

The result section is divided into two parts. The first examines the results from the prediction task, specifically assessing the musical features that led listeners to predict the end of a phrase successfully. The second analyzes the data from the rating task.

A. Prediction Task Results

For this part of the analysis, a participant successfully predicted an ending if their press fell within a 2000 ms window around endings examined in this study. These windows began 500 ms before the onset of the last note of a formal unit and lasted for 1500 ms after this point. If there was a suspension, the window began 500 ms before the beginning of the goal harmony. None of the windows overlapped. Response time was measured from the onset of the last note, so responses made prior to the arrival of the ending received a negative response time while responses occurring after received a positive response time. Musical expertise is indicated by group, which is a three-level variable (non-musicians = 1, undergraduate musicians = 2, and graduate musicians = 3). Cadence and hypermeter are binary, within-subject variables (no cadence, more or less than four or eight measures = 0; presence of a cadence, four or eight measures = 1).

The first set of analyses examines whether participants are more likely to predict the ending of a phrase if it concludes either with a cadence or a temporal distance of four or eight measures from the previous phrase ending (see Table 1). The significant value for group (in the second row) indicates that as musical expertise increases, so does the participant's ability to predict the ends of phrases. Of the two independent within-subject variables, only the presence of a cadence significantly predicts a participant response. Further, there is a significant interaction between musical expertise and the presence of a cadential gesture (see Figure 1). As musical expertise increases, participants are more likely to respond at a cadential gesture. There was no main effect for hypermeter.

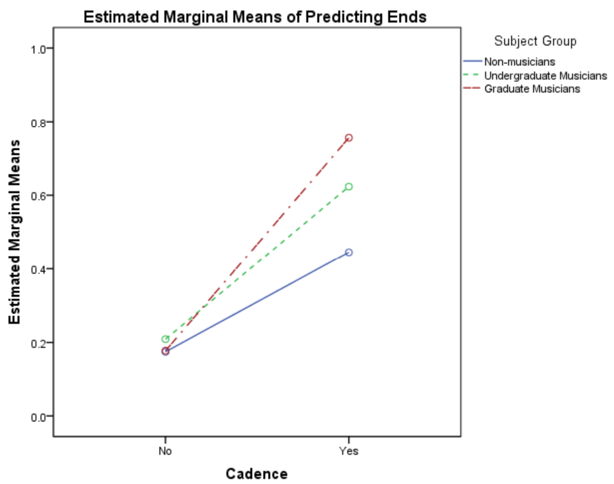

Figure 1. Interaction between subject group and the presence of a cadence.

The cadence variable was then divided into three catagories based on the type of cadence (PAC, HC, and IAC); the results from the mixed logit regression analysis using these independent varibles are located in Table 2. While there is not a significant effect for the HC, there are significant main effects for both the PAC and the IAC. Listeners are more than twice as likely to predict an ending when a tonic chord concludes a phrase.

As before, as musical expertise increases, partcipants are more likely to make their predictions within the two-second window around the cadence types. Figure 2 graphs the ANOVA estimated means for each participant group. Compared with the non-musicians, musicians are more likely to respond to a PAC, while only graduate musicians respond to a HC. All three groups respond to the IAC, but to differing extents.

Since the two-second window around each ending is rather long, the response time data can provides a clearer picture as to the confidence of the participant's prediction. A smaller response time would suggest that the participant was better able to anticipate an ending. Most of the response time data was greater than zero, indicating that most responses occur after the onset of the last note. This could reflect the time it take for a participant to phyically respond to a prediction or the innate difficulty of the prediction task, where participants may be responding retrospectively to an ending point despite directions to predict endings.

To analyze latency, data for endings where a participant did not respond were removed, retaining only the data points where a participant responded within a window. The data analysis of response time only considered the influence of cadences on the timing of the prediction. There is a main effect for musical expertise; as musical expertise increases, response time decreases (note the negative coefficient for group in the second row of Table 3). There is a main effect for all three cadence types on the response time. Here a smaller coefficient corresponds with a faster response time for that cadence. PACs ellicit the fastest response time, followed closely by HCs. Both the PAC and IAC have a subject group interaction, where increased expertise results in a faster response time.

Table 1. Mixed Logit Regression Analysis: Cadence and Hypermeter

Fixed Effect[1]	Coefficient	Standard error	*t*-ratio	*d.f.*	*p*-value	Odds Ratio	Confidence Interval
Intercept	-1.176341	0.274569	-4.284	9023	<0.001	0.308405	(0.180,0.528)
Group	0.462805	0.114176	4.053	72	<0.001	1.588524	(1.265,1.995)
Cadence	0.607246	0.168288	3.608	9023	<0.001	1.835371	(1.320,2.553)
Group	0.801740	0.081178	9.876	9023	<0.001	2.229417	(1.901,2.614)
Hypermeter	-0.114407	0.161683	-0.708	9023	0.479	0.891895	(0.650,1.224)
Group	-0.094700	0.077748	-1.218	9023	0.223	0.909646	(0.781,1.059)

[1] The odds ratio in each row for every within-subject variable shows the odds of a participant's response if the feature is present (with the exception of the first row, which is only needed for the regression equation). The "Group" row for each within-subject variable shows the interaction between musical expertise and the independent variable.

Table 2. Mixed logit regression analysis: cadence types.

Fixed Effect	Coefficient	Standard error	*t*-ratio	*d.f.*	*p*-value	Odds Ratio	Confidence Interval
Intercept	-1.298018	0.323561	-4.012	9022	<0.001	0.273072	(0.145,0.515)
Group	0.551098	0.127846	4.311	72	<0.001	1.735157	(1.345,2.239)
PAC	0.784639	0.197140	3.980	9022	<0.001	2.191616	(1.489,3.225)
Group	1.260553	0.103436	12.187	9022	<0.001	3.527371	(2.880,4.320)
HC	0.109291	0.182932	0.597	9022	0.550	1.115487	(0.779,1.597)
Group	0.708266	0.085718	8.263	9022	<0.001	2.030468	(1.716,2.402)
IAC	0.913797	0.221879	4.118	9022	<0.001	2.493774	(1.614,3.852)
Group	0.420757	0.104090	4.042	9022	<0.001	1.523114	(1.242,1.868)

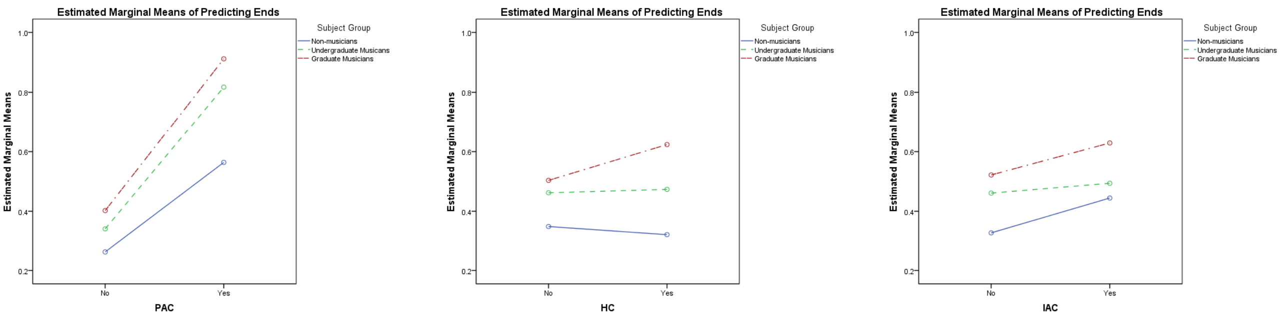

Figure 2. Interactions between subject group and cadence type.

Table 3. Mixed models regression analysis: response time and cadence types.

Fixed Effect	Coefficient	Standard error	*t*-ratio	*d.f.*	*p*-value
Intercept	0.821587	0.072173	11.384	3986	<0.001
Group	-0.125998	0.031354	-4.019	72	<0.001
PAC	0.104156	0.043245	2.408	3986	0.016
Group	-0.071626	0.019831	-3.612	3986	<0.001
HC	0.135055	0.046483	2.905	3986	0.004
Group	-0.035412	0.021012	-1.685	3986	0.092
IAC	0.264896	0.052113	5.083	3986	<0.001
Group	-0.095702	0.023666	-4.044	3986	<0.001

B. Rating Task Results

Along with the between-subject group level, three independent within-subject variables were included in the analysis of the rating data (see Table 4). These were whether the participant anticipated a particular end in the prediction task, the length of the excerpt, and the presence of a cadence. Both the "predicted" and cadence variables are binary variable (1 = participant predicted that particular ending in the previous task and 1 = presence of a cadence) while length is measured in seconds. For each of these variables, a higher value corresponds to a significant increase in the rating of closure. Before taking into account any of the independent variables, there is no significant difference between the ratings made by subjects with different levels of musical expertise, but there are interactions between subject group and predicted ends, as well as subject group and the presence of a cadence. Participants with more musical expertise consistently rate the clips higher when these variables are present.

Table 4. Mixed models regression analysis: ratings.

Fixed Effect	Coefficient	Standard error	*t*-ratio	*d.f.*	*p*-value
Intercept	1.85109	0.310883	5.954	3643	<0.001
Group	0.05496	0.142029	0.387	72	0.700
Predict	0.41812	0.157657	2.652	3643	0.008
Group	0.27069	0.076986	3.516	3643	<0.001
Length	0.28816	0.031036	9.285	3643	<0.001
Group	-0.02238	0.014364	-1.558	3643	0.119
Cadence	1.29699	0.162001	8.006	3643	<0.001
Group	0.19181	0.078961	2.429	3643	0.015

Table 5. Mixed models regression analysis: ratings and response time.

Fixed Effect	Coefficient	Standard error	*t*-ratio	*d.f.*	*p*-value
Intercept	5.04466	0.297488	16.958	1875	<0.001
Group	0.18669	0.128796	1.449	72	0.152
RT	-0.28242	0.311667	-0.906	1875	0.365
Group	-0.40543	0.152425	-2.660	1875	0.008

The final analysis uses only the rating data from the clips in which the participant successfully predicted the ending to see if there is a correlation between ratings and response time. A negative coefficient for response time would have indicated that as the response times increase, the ratings would decrease, but Table 6 shows no main effect for response time. There is, however, an interaction: participants with more musical expertise are more likely to give the clips with a faster response time in the prediction task a higher rating.

IV. DISCUSSION

Overall, the data support the hypothesis that anticipated musical endings evoke a feeling of closure. The data illustrate a correlation between a listener's ability to predict an ending as the composition unfolds and that listener's subsequent rating of closure for that particular ending. This finding supports the argument that the second segmentation principle by Pearce, Müllensiefen, and Wiggins, segmentation occurs before events with a relatively high-level of uncertainty, is more associated with the perception of finality than a boundary solely resulting from a discontinuity in sound.

For participants to be successful in the prediction task, they had to rely on pre-existing knowledge structures that capture stereotypical ending gestures in this style. From a theoretical perspective, these ending gestures are represented by traditional cadences. Cadences concluding with a tonic harmony were better predicted in the first task (Table 2). From the standpoint of probabilistic learning, it would be easier to predict the tonic arrival of the PAC and the IAC since it always follows a root position dominant harmony, resulting in a high transitional probability. In contrast, because the dominant harmony of the HC is not always preceded by the same harmony, a listener may not be able to accurately predict its arrival. Further there are a few instances of cadential ambiguity in these movements, especially in regards to the HC. In his 2010 talk at the Annual Meeting of the Society for Music Theory, Burstein effectively demonstrated that distinguishing between a HC and an elided PAC could be difficult, especially when there is continuous motion from the dominant chord of the HC to the tonic beginning of the next phrase. Measures 54–55 of the G-major Quartet represent one such case of this type of cadential ambiguity (see Figure 3): despite the convincing arrival on the dominant in m. 54, listeners could interpret the cello's downward motion into the tonic pitch on the downbeat of m. 55 as the ending instead. A similar situation occurs in mm. 24–25 of the same quartet (see Figure 4). Here an arrival on the dominant in m. 24 marks the end of the B section, but at this point the second violin initiates a gesture that leads into the return of the A section. It could be possible that without a break in the sound and the high transitional probability that a dominant chord will lead to tonic, listeners would not experience closure in m. 24, but rather a retrospective boundary when the opening theme recurs. This ambiguity surrounding the HC may further explain the lack of a main effect for this cadence type.

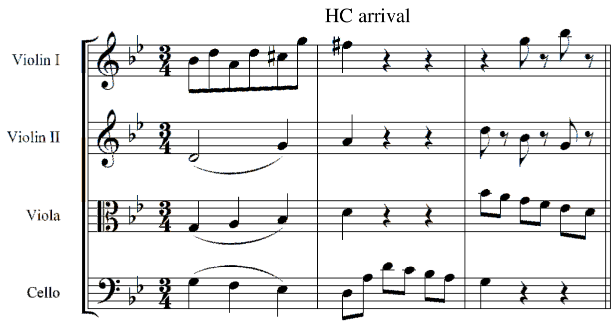

Figure 3. Mozart, String Quartet No. 3 in G Major, K. 156, mm. 53–55.

Figure 4. Mozart, String Quartet No. 3 in G Major, K. 156, mm. 23–25.

Further supporting the feeling of finality associated with cadential arrival, participants responded more quickly to a PAC (generally considered the most final-sounding cadence) than to the other cadences (Table 3). Participants also tended to respond quicker to the HC compared with the IAC. Because this data set only used points where participants responded, it may have removed the more ambiguous HCs, leaving those that were especially predictable.

As seen in previous studies, musical experience influenced the participants' perception of closure (Eberlein & Fricke, 1992; Neuhaus, Knösche, & Friderici, 2006). Main effects were magnified for the participants with more musical experience: participants with more experience successfully predicted more endings, especially those marked by a cadence, better than participants with less experience did. Their ability to anticipate endings more quickly and accurately in the prediction task suggests that the participants with increased musical experience drew from knowledge structures supported by many more exemplars of common ending paradigms in this style. Further, these experienced participants showed a significant correlation between their ratings and their data from the prediction task: both the endings they predicted and their faster response times correlate with higher ratings for closure.

In the rating task, participants tended to rate longer clips as more closed. This length effect has occurred in similar rating tasks when determining the degree of closure of an excerpt from a longer composition (Joichi, 2006). Longer excerpts may evoke a greater feeling of finality because the longer context can form the basis for more accurate predictions. This was one of the few variables that did not have a subject group interaction, suggesting that the effect did not rely on a sophisticated implicit knowledge of stylistic norms.

Event Segmentation Theory posits that an increase in the transient prediction error creates a perceptual boundary. While this study does not directly measure the increase in this transient prediction error following a cadence, it is safe to assume that the ability of a listener to predict upcoming musical material following a cadence is lower than their ability to predict the cadential arrival. The correlation between the prediction and rating tasks further suggests that larger increases in prediction error results in a hierarchically significant musical boundary, eliciting a stronger feeling of closure.

ACKNOWLEDGMENT

This research was supported by a Florida State University Dissertation Research Grant.

REFERENCES

Agawu, V. K. (1987). Concepts of closure and Chopin's opus 28. *Music Theory Spectrum, 9*, 1-17.

Anson-Cartwright, M. (2007). Concepts of closure in tonal music: A critical study. *Theory and Practice, 32*, 1-17.

Burstein, P. (November, 2010). Half, full or in between? Distinguishing between half and authentic cadences. Paper presented at The Annual Meeting of the Society for Music Theory, Indianapolis, IN.

Caplin, W. E. (2004). The classical cadence: Conceptions and misconceptions. *Journal of the American Musicological Society, 57*(1), 51-117.

Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's Grouping Preference Rules. *Music Perception 4*(4), 325-60.

Eberlein, R., & Fricke, J.P. (1992). *Kadenzwahrnegmung und Kadenzgeschichte: Ein Beitrag zu einer Grammatik der Musik*. Frankfurt am Main: Peter Lang.

Hopkins, R. G. (1990). *Closure and Mahler's music: The role of secondary parameters*. Philadelphia: University of Pennsylvania Press.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

Joichi, J. M. (2006). *Closure, context, and hierarchical grouping in music. A theoretical and empirical investigation.* Doctoral dissertation, Northwestern University.

Knösche, T. R., Neuhaus, C. Jens Haueisen, J., Alter, K., Maess, B., Witte, O. W., & Friederici, A. D. (2005). Perception of phrase structure in music. *Human Brian Mapping, 24*, 259-73.

Krumhansl, C. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas." *Music Perception*, *13*(3), 401-32.

Kurby, C. A. & Zacks, J. M. (2007). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences, 12*, 72-79.

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

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

———. (1973). *Explaining music*. Chicago: University of Chicago Press.

Mozart, W.A. (1988). String Quartets Nos. 3, 8, 13. [Recorded by the Amadeus Quartet]. *Mozart: The String Quartets* [CD]. Hamburg: Polydor International GmbH.

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

Neuhaus, C., Knösche, T. R., & Friederici, A. D. (2006). Effects of musical expertise and boundary markers on phrase perception in music. *Journal of Cognitive Neuroscience, 18*(3), 472-93.

Snyder, B. (2000). *Music and memory*. Cambridge: MIT Press.

Zacks, J. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28*, 979-1008.