Simulating Melodic and Harmonic Expectations for Tonal Cadences Using Probabilistic Models

David R. W. Sears

McGill University, Montreal, Canada

Marcus T. Pearce

Queen Mary University of London, UK

William E. Caplin and Stephen McAdams

McGill University, Montreal, Canada

Accepted for publication in *Journal of New Music Research*, 14 July 2017 (in press)

Author Note

Funding was provided by a Richard H. Tomlinson fellowship and a Quebec doctoral fellowship from the Programme de bourses d'excellence pour étudiants étrangers awarded to David Sears, a UK Engineering and Physical Sciences Research Council grant (EP/M000702/1) awarded to Marcus Pearce, a Canadian Social Sciences and Humanities Research Council grant (410-2010-1091) and James McGill Professorship awarded to William E. Caplin, and Canadian Natural Sciences and Engineering Research Council (RGPIN 2015-05280) and Social Sciences and Humanities Research Council (410-2009-2201) grants and Canada Research Chair (950-223484) awarded to Stephen McAdams.

Correspondence: David R. W. Sears, McGill University, Schulich School of Music, Strathcona Music Building, 555 Sherbrooke St. W., Montreal, QC H3A 1E3.

E-mail: david.sears@jku.at

Abstract

This study examines how the mind's predictive mechanisms contribute to the perception of cadential closure during music listening. Using the Information Dynamics of Music model (or IDyOM) to simulate the formation of schematic expectations—a finite-context (or *n*-gram) model that predicts the next event in a musical stimulus by acquiring knowledge through unsupervised statistical learning of sequential structure—we predict the terminal melodic and harmonic events from 245 exemplars of the five most common cadence categories from the classical style. Our findings demonstrate that (1) terminal events from cadential contexts are more predictable than those from non-cadential contexts; (2) models of cadential strength advanced in contemporary cadence typologies reflect the formation of schematic expectations; and (3) a significant decrease in predictability follows the terminal note and chord events of the cadential formula.

Keywords: cadence, expectation, statistical learning, segmental grouping, *n*-gram models

Simulating Melodic and Harmonic Expectations
for Tonal Cadences Using Probabilistic Models

1 Introduction

In the intellectual climate now prevalent, many scholars view the brain as a 'statistical sponge' whose purpose is to predict the future (Clark, 2013). While descending a staircase, for example, even slightly misjudging the height or depth of each step could be fatal, so the brain predicts future steps by building a mental representation of the staircase, using incoming auditory, visual, haptic, and proprioceptive cues to minimize potential prediction errors and update the representation in memory. Researchers sometimes call these representations *schemata*—'active, developing patterns' whose units are serially organized, not simply as individual members coming one after the other, but as a unitary mass (Bartlett, 1932, p. 201). Over the course of exposure, these schematic representations obtain greater specificity, thereby increasing our ability to navigate complex sensory environments and predict future outcomes.

Among music scholars, this view was first crystallized by Meyer (1956, 1967), with the resurgence of associationist theories in the cognitive sciences—which placed the brain's predictive mechanisms at the forefront of contemporary research in music psychology—following soon thereafter. Krumhansl (1990) has suggested, for example, that composers often exploit the brain's potential for prediction by organizing events on the musical surface to reflect the kinds of statistical regularities that listeners will learn and remember. The tonal cadence is a case in point. As a recurrent temporal formula appearing at the ends of phrases, themes, and larger sections in music of the common-practice period, the cadence provides perhaps the clearest instance of phrase-level schematic organization in the tonal system. To be sure, cadential formulæ flourished in eighteenth-century compositional practice by serving to 'mark the breathing places in the music, establish the tonality, and render coherent the formal structure,' thereby cementing their position 'throughout the entire period of common harmonic practice' (Piston, 1962, p.

108). As a consequence, Sears (2015, 2016) has argued that cadences are learned and remembered as *closing schemata*, whereby the initial events of the cadence activate the corresponding schematic representation in memory, allowing listeners to form expectations for the most probable continuations in prospect. The subsequent realization of those expectations then serves to close off both the cadence itself, and perhaps more importantly, the longer phrase-structural process that subsumes it.

There is a good deal of support for the role played by expectation and prediction in the perception of closure (Huron, 2006; Margulis, 2003; Meyer, 1956; Narmour, 1990), with scholars also sometimes suggesting that listeners possess schematic representations for cadences and other recurrent closing patterns (Eberlein & Fricke, 1992; Eberlein, 1997; Gjerdingen, 1988; Meyer, 1967; Rosner & Narmour, 1992; Temperley, 2004). Yet currently very little experimental evidence justifies the links between expectancy, prediction, and the variety of cadences in tonal music or indeed, more specifically, in music of the classical style (Haydn, Mozart, and Beethoven), where the compositional significance of cadential closure is paramount (Caplin, 2004; Hepokoski & Darcy, 2006; Ratner, 1980; Rosen, 1972). This point is somewhat surprising given that the tonal cadence is the quintessential compositional device for suppressing expectations for further continuation (Margulis, 2003). The harmonic progression and melodic contrapuntal motion within the cadential formula elicit very definite expectations concerning the harmony, the melodic scale degree, and the metric position of the goal event. As Huron puts it, ‘it is not simply the final note of the cadence that is predictable; the final note is often approached in a characteristic or formulaic manner. If cadences are truly stereotypic, then this fact should be reflected in measures of predictability’ (2006, p. 154). If Huron is right, applying a probabilistic approach to the cadences from a representative corpus should allow us to examine these claims empirically.

This study applies and extends a probabilistic account of expectancy formation called the Information Dynamics of Music model (or IDyOM)—a finite-context (or *n*-gram) model that predicts the next event in a musical stimulus by acquiring knowledge through unsupervised statistical learning of sequential structure—to examine how the formation, fulfillment, and violation of schematic expectations may contribute to the perception of cadential closure during music listening (Pearce, 2005). IDyOM is based on a class of Markov models commonly used in statistical language modeling (Manning & Schütze, 1999), the goal of which is to simulate the learning mechanisms underlying human cognition. Pearce explains,

> It should be possible to design a statistical learning algorithm ... with no initial knowledge of sequential dependencies between melodic events which, given exposure to a reasonable corpus of music, would exhibit similar patterns of melodic expectation to those observed in experiments with human subjects (Pearce, 2005, p. 152).

Unlike language models, which typically deal with unidimensional inputs, IDyOM generates predictions for multidimensional melodic sequences using the multiple viewpoints framework developed by Conklin (1988, 1990; Conklin & Witten, 1995), which is to say that Pearce's model generates predictions for viewpoints like chromatic pitch by combining predictions from a number of potential viewpoints using a set of simple heuristics to minimize model uncertainty (Pearce, Conklin, & Wiggins, 2005). In the past decade, studies have demonstrated the degree to which IDyOM can simulate the responses of listeners in tasks involving melodic segmentation (Pearce, Müllensiefen, & Wiggins, 2010), subjective ratings of predictive uncertainty (Hansen & Pearce, 2014), subjective and psychophysiological emotional responses to expectancy violations (Egermann, Pearce, Wiggins, & McAdams, 2013), and behavioral (Omigie, Pearce, & Stewart, 2012; Pearce & Wiggins, 2006; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010), electrophysiological (Omigie, Pearce, Williamson, & Stewart, 2013), and neural measures of melodic pitch expectations (Pearce, Ruiz, et al., 2010). And yet, the majority of these studies were limited to the simulation of melodic pitch expectations, so this investigation develops new representation schemes that also permit the probabilistic modeling of harmonic sequences in complex polyphonic textures.

To consider how IDyOM might simulate schematic expectations in cadential contexts, this study adopts a corpus-analytic approach, using the many methods of statistical inference developed in the experimental sciences to examine a few hypotheses about cadential expectancies. To that end, §2 provides a brief summary and discussion of the cadence concept, as well as the typology on which this study is based (Caplin, 1998, 2004), and then offers three hypotheses designed to examine the link between prediction and cadential closure. Next, §3 introduces the multiple viewpoints framework employed by IDyOM, and §4 describes the methods for estimating the conditional probability function for individual melodic or harmonic viewpoints using maximum likelihood (ML) estimation and the prediction-by-partial-match (PPM) algorithm. We then present in §5 the corpus of expositions and the annotated cadence collection from Haydn's string quartets and describe Pearce's procedure for improving model performance by combining viewpoint models into a single composite prediction for each melodic or harmonic event in the sequence. Finally, §6 presents the results of the computational experiments, and §7 concludes by discussing limitations of the modelling approach and considering avenues for future research.

2 The Classical Cadence

Like many of the concepts in circulation in music scholarship (e.g., tonality, harmony, phrase, meter), the cadence concept has been extremely resistant to definition. To sort through the profusion of terms associated with cadence, Blombach (1987) surveyed definitions in eighty-one textbooks distributed around a median publication date of 1970. Her findings suggest that the cadence is most frequently characterized as a *time span*, which consists of a conventionalized harmonic progression, and in some instances, a 'falling' melody. In over half of the textbooks surveyed, these harmonic and melodic formulæ are also classified into a compendium of cadence types, with the degree of finality associated with each type sometimes leading to comparisons with punctuation in language. However, many of these definitions also conceptualize the cadence as a 'point of arrival' (Ratner, 1980), or *time point*, which marks the conclusion of an ongoing phrase-structural process, and which is often characterized as a moment of rest, quiescence, relaxation, or repose. Thus a cadence is simultaneously understood as time-span and time-point, the former relating to its most representative (or recurrent) features (cadence as *formula*), the latter to the presumed boundary it precedes and engenders (cadence as *ending*) (Caplin, 2004).

The compendium of cadences and other conventional closing patterns associated with the classical period is enormous, but contemporary scholars typically cite only a few, which may be classified according to two fundamental types: those cadences for which the goal of the progression is tonic harmony (e.g., perfect authentic, imperfect authentic, deceptive, etc.), and those cadences for which the goal of the progression is dominant harmony (e.g., half cadences). Table 1 provides the harmonic and melodic characteristics for five of the most common cadence categories from Caplin's typology (1998, 2004). The perfect authentic cadence (PAC), which features a harmonic progression from a root-position dominant to a root-position tonic, as well as the arrival of the melody on $\hat{1}$, serves as the quintessential closing pattern not only for the high classical period (Gjerdingen, 2007), but for repertories spanning much of the history of Western music. The imperfect authentic cadence (IAC) is a melodic variant of the PAC category that replaces $\hat{1}$ with $\hat{3}$ (or, more rarely, $\hat{5}$) in the melody, and like the PAC category, typically appears at the conclusion of phrases, themes, or larger sections.

The next two categories represent cadential deviations, in that they initially promise a perfect authentic cadence, yet fundamentally deviate from the pattern's terminal events, thus failing to achieve authentic cadential closure at the expected moment Caplin calls *cadential arrival* (1998, p. 43). The deceptive cadence (DC) leaves harmonic closure somewhat open by closing with a non-tonic harmony, usually vi, but the melodic line resolves to a stable scale degree like $\hat{1}$ or $\hat{3}$, thereby providing a provisional sense of ending for the ongoing thematic process. The evaded cadence is characterized by a sudden interruption in the projected resolution of the cadential process. For example, instead of resolving to $\hat{1}$, the melody often leaps up to some other scale degree like $\hat{5}$, thereby replacing the expected ending with material that clearly initiates the subsequent process. Thus, the evaded cadence projects no sense of ending whatsoever, as the events at the expected moment of cadential arrival, which should group backward by ending the preceding thematic process, instead group forward by initiating the subsequent process. Finally, the half cadence (HC) remains categorically distinct from both the authentic cadence categories and the cadential deviations, since its ultimate harmonic goal is dominant (and not tonic) harmony. The HC category also tends to be defined more flexibly than the other categories in that the terminal harmony may support any chord member in the soprano (i.e., $\hat{2}$, $\hat{5}$, or $\hat{7}$).

This study examines three claims about the link between prediction and cadential closure. First, if cadences serve as the most predictable, probabilistic, specifically envisaged formulæ in all of tonal music (Huron, 2006; Meyer, 1956), we would expect terminal events from cadential contexts to be more predictable than those from non-cadential contexts even if both contexts share similar or even identical terminal events (e.g., tonic harmony in root position, $\hat{1}$ in the melody, etc.). Thus, Experiment 1 examines the hypothesis that cadences are more predictable than their non-cadential counterparts by comparing the probability estimates obtained from IDyOM for the terminal events from the PAC and HC categories—the two most prominent categories in tonal music—with those from non-cadential contexts that share identical terminal events.

Second, applications of cadence typologies like the one employed here often note the correspondence between cadential strength (or finality) on the one hand and expectedness (or predictability) on the other. Dunsby has noted, for example, that in Schoenberg's view, the experience of closure for a given cadential formula is only satisfying to the extent that it fulfills a stylistic expectation (1980, p. 125). This would suggest that the strength and specificity of our schematic expectations formed in prospect and their subsequent realization in retrospect contributes to the perception of cadential strength, where the most expected (i.e., probable) endings are also the most complete or closed. Sears (2015) points out that models of cadential strength advanced in contemporary cadence typologies typically fall into two categories: those that compare every cadence category to the perfect authentic cadence (Latham, 2009; Schmalfeldt, 1992), called the *1-schema* model; and those that distinguish the PAC, IAC, and HC categories from the cadential deviations because the former categories allow listeners to generate expectations as to how they might end, called the *Prospective* (or *Genuine*) *Schemas* model (Sears, 2015, 2016). In the 1-schema model, the half cadence represents the weakest cadential category; it is marked not by a deviation in the melodic and harmonic context at cadential arrival (such as the deceptive or evaded cadences), but rather by the absence of that content, resulting in the following ordering of the cadence categories based on their perceived strength, $PAC \rightarrow IAC \rightarrow DC \rightarrow EV \rightarrow HC$. In the *Prospective Schemas* model, however, the half cadence is a distinct closing schema that allows listeners to generate expectations for its terminal events, and so represents a stronger ending than the aforementioned cadential deviations, resulting in the ordering, $PAC \rightarrow IAC \rightarrow HC \rightarrow DC \rightarrow EV$ (for further details, see Sears, 2015). Experiment 2 directly compares these two models of cadential strength.

Third, a number of studies have supported the role played by predictive mechanisms in the segmentation of temporal experience (Brent, 1999; P. Cohen, Adams, & Heeringa, 2007; Elman, 1990; Kurby & Zacks, 2008; Pearce, Müllensiefen, & Wiggins, 2010; Peebles, 2011). In event segmentation theory (EST), for example, perceivers form working memory representations of ‘what is happening now,’ called *event models*, and discontinuities in the stimulus elicit prediction errors that force the perceptual system to update the model and segment activity into discrete time spans, called *events* (Kurby & Zacks, 2008). In the context of music, such discontinuities can take many forms: sudden changes in melody, harmony, texture, surface activity, rhythmic duration, dynamics, timbre, pitch register, and so on. What is more, when the many parameters effecting segmental grouping act together to produce closure at a particular point in a composition, cadential or otherwise, *parametric congruence* obtains (Meyer, 1973). Thus, Experiment 3 examines whether (1) the terminal event of a cadence, by serving as a predictable point of closure, is the most expected event in the surrounding sequence; and (2) the next event in the sequence, which initiates the subsequent musical process, is comparatively unexpected. Following EST, the hypothesis here is that unexpected events engender prediction errors that lead the perceptual system to segment the event stream into discrete chunks (Kurby & Zacks, 2008). If the terminal events from genuine cadential contexts are highly predictable, then prediction errors for the comparatively unpredictable events that follow should force listeners to segment the preceding cadential material. For the cadential deviations, however, prediction errors should occur at, rather than following, the terminal events of the cadence.

3 Multiple Viewpoints

Most natural languages consist of a finite alphabet of discrete symbols (letters), combinations of which form words, phrases, and so on. As a result, the mapping between the individual letter or word encountered in a printed text and its symbolic representation in a computer database is essentially one-to-one. Music encoding is considerably more complex. Notes, chords, phrases, and the like are characterized by a number of different features, and so regardless of the unit of meaning, digital encodings of individual events must concurrently represent multiple properties of the musical surface. To that end, many symbolic formats employ some variant of the *multiple viewpoints* framework first proposed by Conklin (1988, 1990; Conklin & Witten, 1995), and later extended and refined by Pearce (Pearce & Wiggins, 2004; Pearce, 2005; Pearce et al., 2005).

The multiple viewpoints framework accepts sequences of musical events that typically correspond to individual notes as notated in a score, but which may also include composite events like chords. Each event e consists of a set of *basic attributes*, and each attribute is associated with a *type*, τ, which specifies the properties of that attribute. The *syntactic domain* (or alphabet) of each type, $[\tau]$, denotes the set of all unique elements associated with that type, and each element of the syntactic domain also maps to a corresponding set of elements in the *semantic domain*, $[\![\tau]\!]$. Following Conklin, attribute types appear here in typewriter font to distinguish them from ordinary text. To represent a sequence of pitches as scale degrees derived from the twelve-tone chromatic scale, for example, the type *chromatic scale degree* (or `csd`) would consist of the syntactic set, $\{0, 1, 2, ..., 11\}$, and the semantic set, $\{\hat{1}, \sharp\hat{1}/\flat\hat{2}, \hat{2}, ..., \hat{7}\}$, where 0 represents $\hat{1}$, 7 represents $\hat{5}$, and so on (see Figure 1).

Within this representation language, Conklin and Witten (1995) define several distinct classes of type, but this study examines just three: *basic*, *derived*, and *linked*. Basic types are irreducible representations of the musical surface, which is to say that they cannot be derived from any other type. Thus, an attribute representing the sequence of pitches from the twelve-tone chromatic scale—hereafter referred to as *chromatic pitch*, or `cpitch`—would serve as a basic type in Conklin's approach because it cannot be derived from a sequence of pitch classes, scale degrees, melodic intervals, or indeed, any other attribute. What is more, basic types represent every event in the corpus. For example, a sequence of melodic contours would not constitute a basic type because either the first or last events of the melody would receive no value. Indeed, an interesting property of the set of n basic types for any given corpus is that the Cartesian product of the domains of those types determines the *event space* for the corpus, denoted by ξ:

$$\xi = [\tau_1] \times [\tau_2] \times ... \times [\tau_n]$$

Each event consists of an n-tuple in ξ—a set of values corresponding to the set of basic types that determine the event space. ξ therefore denotes the set of all representable events in the corpus (Pearce, 2005).

As should now be clear from the examples given above, derived types like pitch class, scale degree, and melodic interval do not appear in the event space but are derived from one or more of the basic types. Thus, for every type in the encoded representation there exists a partial function, denoted by Ψ, which maps sequences of events onto elements of type τ. The term *viewpoint* therefore refers to the function associated with its type, but for convenience Conklin and Pearce refer to viewpoints by the types they model.[1] The function is partial because the output may be undefined for certain events in the sequence (denoted by $\perp$). Again, viewpoints for attributes like melodic contour or melodic interval demonstrate this point, since either the first or last element will receive no value (i.e., it will be undefined).

Basic and derived types attempt to model the relations within attributes, but they fail to represent the relations *between* attributes. Prototypical utterances like cadences, for example, are necessarily comprised of a cluster of co-occurring features, so it is important to note that the relations between those features could be just as significant as their presence (or absence) (Gjerdingen, 1991). This is to say that the harmonic progression V–I presented in isolation does not provide sufficient grounds for the identification of a perfect authentic cadence, but the co-occurrence of that progression with $\hat{1}$ in the soprano, a six-four sonority preceding the root-position dominant, or a trill above the dominant makes such an interpretation far more likely. Linked viewpoints attempt to model correlations between these sorts of attributes by calculating the cross-product of their constituent types.

4 Finite-Context Models

4.1 Maximum Likelihood Estimation

The goal of finite-context models like IDyOM is to derive from a corpus of example sequences a model that estimates the probability of event e_i given a preceding sequence of events e_1 to e_{i-1}, notated here as e_1^{i-1}. Thus, the function $p(e_i|e_1^{i-1})$ assumes that the identity of each event in the sequence depends only on the events that precede it. In principle, the length of the context is limited only by the length of the sequence e_1^{i-1}, but context models typically stipulate a global order bound such that the probability of the next event depends only on the previous $n-1$ events, or $p(e_i|e_{(i-n)+1}^{i-1})$. Following the Markov assumption, the model described here is an $(n-1)^{\text{th}}$ order Markov model, but researchers also sometimes call it an n-gram model because the sequence $e_{(i-n)+1}^{i}$ is an n-gram consisting of a *context* $e_{(i-n)+1}^{i-1}$ and a single-event *prediction* e_i.

To estimate the conditional probability function $p(e_i|e_{(i-n)+1}^{i-1})$ for each event in the *test* sequence, IDyOM first acquires the frequency counts for a collection of such sequences from a *training* set. When the trained model is exposed to the test sequence, it then uses the frequency counts to estimate the probability distribution governing the identity of the next event in the sequence given the $n-1$ preceding events (Pearce, 2005). In this case, IDyOM relies on *maximum likelihood* (ML) estimation.

$$p(e_i|e_{(i-n)+1}^{i-1}) = \frac{c(e_i|e_{(i-n)+1}^{i-1})}{\sum_{e \in A} c(e|e_{(i-n)+1}^{i-1})} \tag{1}$$

The numerator terms represent the frequency count c for the n-gram $e_i|e_{(i-n)+1}^{i-1}$, and the denominator terms represent the sum of the frequency counts c associated with all of the possible events e in the alphabet A following the context $e_{(i-n)+1}^{i-1}$.

4.2 Performance Metrics

To evaluate model performance, the most common metrics derive from information-theoretic measures introduced by Shannon (1948, 1951). Returning to Equation 1, if the probability of e_i is given by the conditional probability function $p(e_i|e_{(i-n)+1}^{i-1})$, *information content* (IC) represents the minimum number of bits required to encode e_i in context (MacKay, 2003).

$$\text{IC}(e_i|e_{(i-n)+1}^{i-1}) = \log_2 \frac{1}{p(e_i|e_{(i-n)+1}^{i-1})} \tag{2}$$

IC is inversely proportional to p and so represents the degree of contextual *unexpectedness* or surprise associated with e_i. Researchers often prefer to report IC over p because it has a more convenient scale (p can become vanishingly small), and since it also has a well-defined interpretation in data compression theory (Pearce, Ruiz, et al., 2010), we will prefer it in the analyses that follow.

Whereas IC represents the degree of unexpectedness associated with a particular event e_i in the sequence, *Shannon entropy* (H) represents the degree of contextual uncertainty associated with the probability distribution governing that outcome, where the probability estimates are independent and sum to one.

$$H(e_{(i-n)+1}^{i-1}) = \sum_{e \in A} p(e_i | e_{(i-n)+1}^{i-1}) \text{IC}(e_i | e_{(i-n)+1}^{i-1}) \tag{3}$$

H is computed by averaging the information content over all e in A following the context $e_{(i-n)+1}^{i-1}$. According to Shannon's equation, if the probability of a given outcome is 1, the probabilities for all of the remaining outcomes will be 0, and $H = 0$ (i.e., maximum certainty). If all of the outcomes are equally likely, however, H will be maximum (i.e., maximum uncertainty). Thus, one can assume that the best performing models will minimize uncertainty.

In practice, we rarely know the true probability distribution of the stochastic process (Pearce & Wiggins, 2004), so it is often necessary to evaluate model performance using an alternative measure called *cross entropy*, denoted by H_m.

$$H_m(p_m, e_1^j) = -\frac{1}{j} \sum_{i=1}^{j} \log_2 p_m(e_i | e_1^{i-1}) \tag{4}$$

Whereas H represents the average information content over all e in the alphabet A, H_m represents the average information content for the model probabilities estimated by p_m over all e in the sequence e_1^j. That is, cross entropy provides an estimate of how uncertain a model is, on average, when predicting a given *sequence* of events (Manning & Schütze, 1999; Pearce & Wiggins, 2004). As a consequence, H_m is often used to evaluate the performance of context models for tasks like speech recognition, machine translation, and spelling correction because, as Brown and his co-authors put it, 'models for which the cross entropy is lower lead directly to better performance' (Brown, Della Pietra, Della Pietra, Lai, & Mercer, 1992, p. 39).

4.3 Prediction by Partial Match (PPM)

Because the number of potential patterns decreases dramatically as the value of n increases, high-order models often suffer from the *zero-frequency problem*, in which n-grams encountered in the test set do not appear in the training set (Witten & Bell, 1991). To resolve this issue, IDyOM applies a data compression scheme called *Prediction by Partial Match* (PPM), which adjusts the ML estimate for each event in the sequence by combining (or *smoothing*) predictions generated at higher orders with less sparsely estimated predictions from lower orders (Cleary & Witten, 1984). Context models estimated with the PPM scheme typically use a procedure called *backoff smoothing* (or *blending*), which assigns some portion of the probability mass from each distribution to an *escape probability* using an escape method to accommodate predictions that do not appear in the training set. When a given event does not appear in the $n-1$ order distribution, PPM stores the escape probability and then iteratively backs off to lower-order distributions until it predicts the event or reaches the zeroth-order distribution, at which point it transmits the probability estimate for a uniform distribution over A (i.e., where every event in the alphabet is equally likely). PPM then multiplies these probability estimates together to obtain the final (smoothed) estimate.

Unfortunately there is no sound theoretical basis for choosing the appropriate escape method (Witten & Bell, 1991), but two recent studies have demonstrated the potential of Moffat's (1990) *method C* to minimize model uncertainty in melodic and harmonic prediction tasks (Hedges, 2016; Pearce & Wiggins, 2004), so we employ that method here.

$$\gamma(e_{(i-n)+1}^{i-1}) = \frac{t(e_{(i-n)+1}^{i-1})}{\sum_{e \in A} c(e|e_{(i-n)+1}^{i-1}) + t(e_{(i-n)+1}^{i-1})} \tag{5}$$

Escape method C represents the escape count t as the number of distinct symbols that follow the context $e_{(i-n)+1}^{i-1}$. To calculate the escape probability for events that do not appear in the training set, γ represents the ratio of the escape count t to the sum of the frequency counts c and t for the context $e_{(i-n)+1}^{i-1}$. The appeal of this escape method is that it assigns greater weighting to higher-order predictions (which are more specific to the context) over lower-order predictions (which are more general) in the final probability estimate (Bunton, 1996; Pearce, 2005). Thus, Equation 1 can be revised in the following way:

$$\alpha(e_i|e_{(i-n)+1}^{i-1}) = \frac{c(e_i|e_{(i-n)+1}^{i-1})}{\sum_{e \in A} c(e|e_{(i-n)+1}^{i-1}) + t(e_{(i-n)+1}^{i-1})} \tag{6}$$

The PPM scheme just described remains the canonical method in many context models (Cleary & Teahan, 1997), but Bunton (1997) has since provided a variant smoothing technique called *mixtures* that generally improves model performance, but which, following Chen and Goodman (1999), we refer to as *interpolated smoothing* (Pearce & Wiggins, 2004). The central idea behind interpolated smoothing is to compute a weighted combination of higher-order and lower-order models for *every* event in the sequence—regardless of whether that event features n-grams with non-zero counts—under the assumption that the addition of lower-order models might generate more accurate probability estimates.[2]

Formally, interpolated smoothing estimates the probability function $p(e_i|e_{(i-n)+1}^{i-1})$ by recursively computing a weighted combination of the $(n-1)^{\text{th}}$ order distribution with the

$(n-2)^{\text{th}}$ order distribution (Pearce & Wiggins, 2004; Pearce, 2005).

$$p(e_i|e_{(i-n)+1}^{i-1}) = \begin{cases} \alpha(e_i|e_{(i-n)+1}^{i-1}) + \gamma(e_{(i-n)+1}^{i-1})p(e_{(i-n)+2}^{i-1}) & \text{if } e_{(i-n)+2}^{i-1} \neq \varepsilon \\ \frac{1}{|A|+1-t(\varepsilon)} & \text{otherwise} \end{cases} \quad (7)$$

In the context of interpolated smoothing, it can be helpful to think of γ as a weighting function, with α serving as the weighted ML estimate. Unlike the backoff smoothing procedure, which terminates at the first non-zero prediction, interpolated smoothing recursively adjusts the probability estimate for each order—regardless of whether the corresponding n-gram features a non-zero count—and then terminates with the probability estimate for ε, which represents a uniform distribution over $|A| + 1 - t(\varepsilon)$ events (i.e., where every event in the alphabet is equally likely). Also note here that in the PPM scheme, the alphabet A increases by one event to accommodate the escape count t but decreases by the number of events in A that never appear in the corpus.[3]

4.4 Variable Orders

The optimal order for context models depends on the nature of the corpus, which in the absence of a priori knowledge can only be determined empirically (2004, p. 2). To resolve this issue, IDyOM employs an extension to PPM called PPM* (Cleary & Teahan, 1997), which includes contexts of variable length and thus 'eliminates the need to impose an arbitrary order bound' (Pearce & Wiggins, 2004, p. 6). In the PPM* scheme, the context length is allowed to vary for each event in the sequence, with the maximum context length selected using simple heuristics to minimize model uncertainty. Specifically, PPM* exploits the fact that the observed frequency of novel events is much lower than expected for contexts that feature exactly one prediction, called *deterministic* contexts. As a result, the entropy of the distributions estimated at or below deterministic contexts tends to be lower than in non-deterministic contexts. Thus, PPM* selects the shortest deterministic context to serve as the global order bound for each event in the sequence. If such a context does not exist, PPM* then selects the longest matching context.

5 Methods

5.1 The Corpus

The corpus consists of symbolic representations of 50 sonata-form expositions selected from Haydn's string quartets (1771–1803). Table 2 presents the reference information, keys, time signatures, and tempo markings for each movement. The corpus spans much of Haydn's mature compositional style (Opp. 17–76), with the majority of the expositions selected from first movements (28) or finales (11), and with the remainder appearing in inner movements (ii: 8; iii: 3). All movements were downloaded from the KernScores database in MIDI format.[4] To ensure that each instrumental part would qualify as monophonic—a pre-requisite for the analytical techniques that follow—all trills, extended string techniques, and other ornaments were removed. For events presenting extended string techniques (e.g., double or triple stops), note events in each part were retained that preserved the voice leading both within and between instrumental parts. Table 3 provides a few descriptives concerning the number of note and chord events in each movement.

To examine model predictions for the cadences in the corpus, we classified exemplars of the five cadence categories that achieve (or at least promise) cadential arrival in Caplin's cadence typology—PAC, IAC, HC, DC, and EV (see Table 1). The corpus contains 270 cadences, but 15 cadences were excluded because either the cadential bass or soprano does not appear in the cello and first violin parts, respectively. Additionally, another 10 cadences were excluded because they imply more than one category (i.e., PAC-EV or DC-EV). Thus, for the analyses that follow, the cadence collection consists of 245 cadences.

Shown in the right-most column of Table 1, the perfect authentic cadence and the half cadence represent the most prevalent categories, followed by the cadential deviations: the deceptive and evaded categories. The imperfect authentic cadence is the least common category, which perhaps reflects the late-century stylistic preference for perfect authentic cadential closure at the ends of themes and larger sections. This distribution also largely replicates previous findings for Mozart's keyboard sonatas (Rohrmeier & Neuwirth, 2011), so it is possible that this distribution may characterize the classical style in general.

5.2 Viewpoint Selection

To select the appropriate viewpoints for the prediction of cadences in Haydn's string quartets, we have adopted Gjerdingen's schema-theoretic approach (2007), which represents the 'core' events of the cadence by the scale degrees and melodic contours of the outer voices (i.e., the two-voice framework), a coefficient representing the strength of the metric position (strong, weak), and a sonority, presented using figured bass notation. Given the importance of melodic intervals in studies of recognition memory for melodies (Dowling, 1981) we might also add this attribute to Gjerdingen's list. However, for the majority of the encoded cadences from the cadence collection, the terminal events at the moment of cadential arrival appear in strong metric positions, and few of the cadences feature unexpected durations or inter-onset intervals at the cadential arrival, so we have excluded viewpoint models for rhythmic or metric attributes from the present investigation, concentrating instead on those viewpoints representing pitch-based (melodic or harmonic) expectations. What is more, IDyOM was designed to combine melodic predictions from two or more viewpoints by mapping the probability distributions over their respective alphabets back into distributions over a single basic viewpoint, such as the pitches of the twelve-tone chromatic scale (i.e., `cpitch`). Thus, for the purposes of model comparison it will also be useful to include `cpitch` as a baseline melodic model in the analyses that follow.

5.2.1 Note Events. Four viewpoints were initially selected to represent note events in the outer parts: chromatic pitch (`cpitch`), melodic pitch interval (`melint`), melodic contour (`contour`), and chromatic scale degree (`csd`). As described previously, `cpitch` represents pitches as integers from 0–127 (in the MIDI representation, C_4 is 60), and serves as the baseline model for the other melodic viewpoint models examined in this study. To derive sequences of melodic intervals, `melint` computes the numerical difference between adjacent events in `cpitch`, where ascending intervals are positive and descending intervals are negative. The viewpoint `contour` then reduces the information present in `melint`, with all ascending intervals receiving a value of 1, all descending intervals a value of -1, and all lateral motion a value of 0. Finally, to relate `cpitch` to a referential tonic pitch class for every event in the corpus, we manually annotated the key, mode, modulations, and pivot boundaries for each movement and then included the analysis in a separate text file to accompany the MIDI representation, both of which appear in the Supplementary materials for each movement in the corpus. Thus, every note event was associated with the viewpoints `key` and `mode`. The vector of keys assumes values in the set $\{0,1,2,...,11\}$, where 0 represents the key of C, 1 represents C♯ or D♭, and so on. Passages in the major and minor modes receive values of 0 and 1, respectively. The viewpoint `csd` then maps `cpitch` to `key` and reduces the resulting vector of chromatic scale degrees modulo 12 such that 0 denotes the tonic scale degree, 7 the dominant scale degree, and so on. By way of example, Figure 1 presents the viewpoint representation for the first violin part from the opening two measures of the first movement of Haydn's String Quartet in E, Op. 17/1.

[Insert Figure 1 about here.]

As mentioned previously, IDyOM is capable of individually predicting any one of these viewpoints using the PPM* scheme, but it can also *combine* viewpoint models for note-event predictions of the same basic viewpoint (i.e., `cpitch`) using a weighted multiplicative combination scheme that assigns greater weights to viewpoints whose predictions are associated with lower entropy at that point in the sequence (Pearce et al., 2005). To determine the combined probability distribution for each event in the test sequence, IDyOM then computes the product of the weighted probability estimates from each viewpoint model for each possible value of the predicted viewpoint.

Furthermore, IDyOM can automate the viewpoint selection process using a hill-climbing procedure called *forward stepwise selection*, which picks the combination of viewpoints that yields the richest structural representations of the musical surface and minimizes model uncertainty. Given an empty set of viewpoints, the stepwise selection algorithm iteratively selects the viewpoint model additions or deletions that yield the most improvement in cross entropy, terminating when no addition or deletion yields an improvement (Pearce, 2005; Potter, Wiggins, & Pearce, 2007). To derive the optimal viewpoint system for the representation of melodic expectations, we employed stepwise selection for the following viewpoints: `cpitch`, `melint`, `csd`, and `contour`. In this case, IDyOM begins with the above set of viewpoint models, but also includes the linked viewpoints derived from that set (i.e., `cpitch` ⊗ `melint`, `cpitch` ⊗ `csd`, `cpitch` ⊗ `contour`, `melint` ⊗ `csd`, `melint` ⊗ `contour`, `csd` ⊗ `contour`), resulting in a pool of ten individual viewpoint models from which to derive the optimal combination of viewpoints.

Viewpoint selection derived the same combination of viewpoint models for the first violin and the cello. For this corpus, `melint` was the best performing viewpoint model in the first step, receiving a cross entropy estimate of 3.006 in the first violin and 2.798 in the cello. In the second step, the combination of `melint` with the linked viewpoint `csd` ⊗ `cpitch` decreased the cross entropy estimate to 2.765 in the first violin and 2.556 in the cello. Including any of the remaining viewpoints did not improve model performance, so the stepwise selection procedure terminated with this combination of viewpoints. In §6, we refer to this viewpoint model as `selection`. What is more, the `contour` model received a much higher cross entropy estimate than the other viewpoint models, so we elected to exclude it in the experiments reported here. Thus, the final melodic viewpoint models selected for the present study are `cpitch`, `melint`, `csd`, and `selection`.

5.2.2 Chord Events. To accommodate chord events, we have extended the multiple viewpoints framework by performing a *full expansion* of the symbolic encoding, which duplicates overlapping note events across the instrumental parts at every unique onset time (Conklin, 2002). This representation yielded two harmonic viewpoints: vertical interval class combination (`vintcc`) and chromatic scale-degree combination (`csdc`). The viewpoint `vintcc` produces a sequence of chords that have analogues in figured-bass nomenclature by modelling the vertical intervals in semitones modulo 12 between the lowest instrumental part and the upper parts from `cpitch`. Unfortunately, however, the syntactic domain of `vintcc` is rather large; the domain of each vertical interval class between any two instrumental parts is $\{0, 1, 2, ..., 11, \perp\}$, yielding 13 possible classes, so the number of combinatorial possibilities for combinations containing two, three, or four instrumental parts is $13^3 - 1$, or 2196 combinations.

To reduce the syntactic domain while retaining those chord combinations that approximate figured bass symbols, Quinn (2010) assumed that the precise location and repeated appearance of a given interval in the instrumental texture are inconsequential to the identity of the combination. Adopting that approach here, we have excluded note events in the upper parts that double the lowest instrumental part at the unison or octave, allowed permutations between vertical intervals, and excluded interval repetitions. As a consequence, the first two criteria reduce the major triads $\langle 4, 7, 0\rangle$ and $\langle 7, 4, 0\rangle$ to $\langle 4, 7, \perp\rangle$, while the third criterion reduces the chords $\langle 4, 4, 10\rangle$ and $\langle 4, 10, 10\rangle$ to $\langle 4, 10, \perp\rangle$. This procedure dramatically reduces the potential domain of `vintcc` from 2196 to 232 unique vertical interval class combinations, though the corpus only contained 190 of the 232 possible combinations, reducing the domain yet further.

To relate each combination to an underlying tonic, the viewpoint `csdc` represents vertical sonorities as combinations of chromatic scale degrees that are intended to approximate Roman numerals. The viewpoint `csdc` includes the chromatic scale degrees derived from `csd` as combinations of two, three, or four instrumental parts. Here, the number of possibilities increases exponentially to $13^4 - 13^1$, or $28, 548$ combinations, since the cello part is now encoded explicitly in combinations containing all four parts. Rather than treating permutable combinations as equivalent (e.g., $\langle 0, 4, 7, \perp\rangle$ and $\langle 4, 7, 0, \perp\rangle$), as was done for `vintcc`, it will also be useful to retain the chromatic scale degree in the lowest instrumental part in `csdc` and only permit permutations in the upper parts. Excluding voice doublings and permitting permutations in the upper parts reduces the potential domain of `csdc` to 2784, though in the corpus the domain reduced yet further to 688 distinct combinations.

Finally, a composite viewpoint was also created to represent those viewpoint models characterizing pitch-based (i.e., melodic *and* harmonic) expectations more generally. To simulate the cognitive mechanisms underlying melodic segmentation, Pearce, Müllensiefen, and Wiggins (2010) found it beneficial to combine viewpoint predictions for basic attributes like chromatic pitch, inter-onset interval, and offset-to-onset interval by multiplying the component probabilities to reach an overall probability for each note in the sequence as the joint probability of the individual basic attributes being predicted. Following their approach, the viewpoint model `composite` represents the product of the `selection` viewpoint model from the first violin (to represent melodic expectations) and the `csdc` viewpoint model (to represent harmonic expectations) for each unique onset time for which a note *and* chord event appear in the corpus. In this case, `csdc` was preferred to `vintcc` in the `composite` model because the former viewpoint explicitly encodes the chromatic scale-degree successions in the lowest instrumental part along with the relevant scale degrees from the upper parts.

5.3 Long-term vs. Short-term

To improve model performance, IDyOM separately estimates and then combines two subordinate models trained on different subsets of the corpus for each viewpoint: a *long-term* model (LTM), which is trained on the entire corpus to simulate long-term, schematic knowledge; and a *short-term* model (STM), which is initially empty for each individual composition and then is trained incrementally to simulate short-term, dynamic knowledge (Pearce & Wiggins, 2012). As a result, the long-term model reflects inter-opus statistics from a large corpus of compositions, whereas the short-term model only reflects intra-opus statistics, some of which may be specific to that composition (Conklin & Witten, 1995; Pearce & Wiggins, 2004). Like the STM, the LTM may also be slightly improved by incrementally training on the composition being predicted, called LTM+. However, the STM only discards statistics when it reaches the end of the composition, so it far surpasses the supposed upper limits for short-term and working memory of around 10–12 s (Snyder, 2000), sometimes by several minutes. What is more, the STM should be irrelevant for the present purposes, since cadences exemplify the kinds of inter-opus patterns that listeners are likely to store in long-term memory. Thus, we have elected to omit the STM in the analyses that follow and only present the probability estimates from LTM+.

5.4 Performance Evaluation

Context models like IDyOM depend on a training set and a test set, but in this case the corpus will need to serve as both. To accommodate small corpora like this one, IDyOM employs a resampling approach called *k-fold cross-validation* (Dietterich, 1998), using cross entropy as a measure of performance (Conklin & Witten, 1995). The corpus is divided into k disjoint subsets containing the same number of compositions, and the LTM+ is trained k times on $k - 1$ subsets, each time leaving out a different subset for testing. IDyOM then computes an average of the k cross entropy values as a measure of the model's performance. Following Pearce and Wiggins (2004), we use 10-fold cross validation for the models that follow.

6 Computational Experiments

6.1 Experiment 1

The perfect authentic and half cadence categories account for 206 of the 245 cadences from the collection, so it seems reasonable that listeners with sufficient exposure to music of the classical style will form schematic expectations for the terminal events of exemplars from these two categories. What is more, if cadences are the most predictable formulæ in all of tonal music, we should expect to find lower IC estimates for the terminal events from the aforementioned cadence categories compared to those from non-cadential closing contexts even if they both share similar or even identical terminal events. Thus, Experiment 1 examines the hypothesis that cadences are more predictable than their non-cadential counterparts.

6.1.1 Analysis. To compare the PAC and HC categories against non-cadential contexts exhibiting varying degrees of closure or stability, each of the viewpoints estimated by IDyOM was analyzed for the terminal note events from the first violin and cello—represented by the viewpoints `cpitch`, `melint`, `csd`, and `selection`—and the terminal chord events from the entire texture—represented by the viewpoints `vintcc`, `csdc`, and `composite`—using a one-way analysis of variance (ANOVA) with a three-level between-groups factor called *closure*. To examine the IC estimates for the first (tonic) type, *tonic closure* consists of three levels: *PAC*, which consists of the IC estimates for the terminal events from the 122 exemplars of the PAC category; *tonic*, which consists of an equal-sized sample of events selected randomly from the corpus that appear in strong metric positions (i.e., appearing at the tactus level; see Sears, 2016) and feature tonic harmony in root position and any scale degree in the soprano; and *non-tonic*, which again consists of an equal-sized sample of events selected randomly from the corpus that appear in strong metric positions, but that feature any other harmony and any other scale degree in the soprano.

To examine the IC estimates for the second (dominant) type, *dominant closure* was designed in much the same way. *HC* consists of the IC estimates for the terminal events from the 84 exemplars of the HC category, while the other two levels consist of equal-sized samples of non-cadential events selected at random. Events from *dominant* appear in strong metric positions and feature dominant harmony in root position with any scale degree in the soprano, while events from *other* appear in strong metric positions but exclude events featuring tonic or dominant harmonies in root position. The assumption behind this additional exclusion criterion is that tonic events in root position are potentially more predictable than root-position dominants in half-cadential contexts (an assumption examined in greater detail in Experiment 2), so it was necessary here to provide a condition that allows us to compare the IC estimates for the terminal events in half-cadential contexts against those featuring other, presumably less stable harmonies and scale degrees.

For every between-groups factor examined in the experiments reported here, Levene's equality of variances revealed significant differences between groups for nearly every viewpoint model. Thus, we employ an alternative to Fisher's F ratio that is generally robust to heteroscedastic data, called the Welch F ratio (Welch, 1951). To determine the effect size both for the Welch F ratio and for the planned comparisons described shortly, we use Cohen's (2008) recent notation of a common effect size measure called estimated ω^2.

To address more specific hypotheses about the potential differences in the IC estimates for the terminal events from cadential and non-cadential contexts, each model also includes two planned comparisons that do not assume equal variances: the first to determine whether the IC estimates from the corresponding cadence category differ significantly from the two non-cadential levels (*Cadences* vs. *Non-Cadences*), and the second to determine whether the IC estimates from the corresponding cadence category differ significantly from the second (tonic or dominant) level of *closure* (*PAC* vs *Tonic* or *HC* vs. *Dominant*). Unfortunately, these additional tests increase the risk of committing a Type I error, so we apply Bonferroni correction to the planned comparisons.

6.1.2 Results. The top bar plots in Figure 2 display the mean IC estimates for the terminal note event in the first violin (left) and cello (right) for each level of *tonic closure*. Table 4 presents the omnibus statistics and planned comparisons. Beginning with the first violin, one-way ANOVAs of the IC estimates revealed a main effect for the viewpoints `melint`, `csd`, and the optimized combination `selection`, but the baseline viewpoint, `cpitch`, was not significant. Mean IC estimates also increased significantly from *PAC* to the non-cadential levels of *tonic closure* for `melint`, `csd`, and `selection`.

Although this trend also emerged for the second planned comparison between *PAC* and *tonic*, only the `melint` model revealed a significant effect. Thus, the viewpoint models for the first violin demonstrated that terminal note events from cadential contexts are more predictable than those from non-cadential contexts.

[Insert Figure 2 about here.]

For the cello, one-way ANOVAs revealed a main effect of *tonic closure* for every viewpoint, but the direction of the effect was reversed. Mean IC estimates decreased in every model from *PAC* to the non-cadential levels of *tonic closure*, as well as from *PAC* to *tonic*. Thus, contrary to our predictions, the terminal events in the cello from cadential contexts were actually *less* predictable than those from non-cadential contexts.

The bottom-left bar plot in Figure 2 displays the mean IC estimates for the terminal chord event—represented by `vintcc` and `csdc`—for each level of the between-groups factor. One-way ANOVAs again revealed a main effect of *tonic closure* for `vintcc` and `csdc`, with the mean IC estimates increasing from *PAC* to the non-cadential levels of *tonic closure*. The second planned comparison comparing *PAC* to *tonic* was not significant for either viewpoint model, however. Thus, for both models the terminal chord events from cadential contexts were more predictable than those from non-cadential contexts.

To represent the predictability of the harmony and melody in a single IC estimate for each note/chord event, we created a `composite` viewpoint that reflects the joint probability of `csdc` and $\texttt{selection}_{\text{vl1}}$. The bottom-right line plot in Figure 2 displays the mean IC estimates for the terminal `composite` event for each level of *tonic closure*. In this case, the one-way ANOVA revealed a significant main effect, with the mean IC estimates increasing from *PAC* to the non-cadential levels of *closure*, and the increase from *PAC* to *tonic* was marginally significant. As a result, `composite` demonstrated an ascending staircase for the levels of *tonic closure*, with *PAC* receiving the lowest IC estimates and *non-tonic* receiving the highest IC estimates.

The top bar plots in Figure 3 display the mean IC estimates for the terminal note event in the first violin (left) and cello (right) for each level of *dominant closure*. For the first violin, one-way ANOVAs revealed a main effect for every viewpoint model. As expected, the mean IC estimates also increased significantly from *HC* to the non-cadential levels of *dominant closure* for every model. The increase from *HC* to *dominant* was also significant for `cpitch`, `melint`, and `selection`, but not for `csd`.

[Insert Figure 3 about here.]

For the cello, the mean IC estimates demonstrated a significant effect of *dominant closure* for `csd`, but the other viewpoint models were not significant. For `csd`, the mean IC estimates increased significantly from *HC* to the non-cadential levels. A similar trend also emerged for the second planned comparison between *HC* and *dominant*, but the effect was not significant. For the viewpoint models representing harmonic progressions, the mean IC estimates revealed main effects of *dominant closure* for `vintcc` and `csdc`, suggesting that the terminal note and chord events represented in the cello and the entire multi-voiced texture are more predictable in half-cadential contexts than in non-cadential contexts. The first planned comparison comparing *HC* with the two non-cadential levels was not significant for these viewpoint models, however. Finally, the `composite` viewpoint demonstrated a significant main effect of *dominant closure*, with the mean IC estimates increasing significantly from *HC* to the non-cadential levels, but not from *HC* to *dominant*.

6.1.3 Discussion. Both between-groups factors demonstrated significantly lower mean IC estimates for the terminal events from cadential contexts compared to those from non-cadential contexts. The factor *tonic closure* elicited significant effects for viewpoints representing both voices of the two-voice framework, with greater effect sizes appearing for viewpoint models characterizing harmonic progressions (`vintcc`, `csdc`, and `composite`). For viewpoint models representing the cello explicitly, however, the terminal events from perfect authentic cadential contexts were actually *less* predictable than those from non-cadential tonic contexts. This finding may reflect limitations of the modelling approach (see §7), but since the leap in the bass by descending fifth (or ascending fourth) in perfect authentic cadential contexts occurs *less frequently* than motion by smaller intervals in any other context (e.g., by unison, m2, or M2) (Sears, 2016), it may also be that cadential bass lines are simply less predictable than their stepwise, non-cadential counterparts when considered in isolation. For the viewpoints that explicitly model the *interaction* between the bass and the upper voices, however (e.g., `vintcc`, `csdc`, or `composite`), IDyOM produced considerably lower IC estimates for cadential successions like $\hat{5}$-$\hat{1}$ than for non-cadential successions like $\hat{1}$-$\hat{1}$, $\hat{2}$-$\hat{1}$, or $\hat{7}$-$\hat{1}$.

For *dominant closure*, significant effects were generally limited to the viewpoint models for `csd` in the outer parts, but the effects were more pronounced for the `csdc` and `composite` models. In each case, the terminal events from cadential contexts were more predictable than those from non-cadential contexts. Nevertheless, half-cadential contexts generally failed to elicit lower mean IC estimates compared to non-cadential root-position dominants. Thus, according to IDyOM, the terminal events from the *HC* level are no more (or less) predictable than any other instance of root-position dominant harmony selected at random from the corpus.

Given our earlier assumptions about schematic expectations for dominant events, these results should not be surprising. Nevertheless, it remains unclear whether terminal events from half cadences receive higher IC estimates on average due to limitations of the modelling approach, because the preceding context fails to stimulate strong expectations for any particular continuation, or because the actual continuation is unexpected (Pearce, Müllensiefen, & Wiggins, 2010, pp. 1374–1375) And yet, by only considering the potential differences between cadential and non-cadential contexts, the previous analysis failed to directly compare the cadence categories from Caplin's typology. We might hypothesize, for example, that the strength and specificity of our schematic expectations formed in prospect and their subsequent realization in retrospect contributes to the perception of cadential *strength*, where the most expected (i.e., probable) endings are also the most complete or closed. From this point of view, the probabilities estimated by IDyOM might correspond with models of cadential strength advanced in contemporary cadence typologies.

6.2 Experiment 2

Recall that the cadence collection consists of exemplars from five categories in Caplin's typology: PAC, IAC, HC, DC, and EV. Sears (2015) recently classified models that estimate the closural strength of these categories into two general types: those that relate every cadential category to one essential prototype, called the *1-Schema* model (Latham, 2009; Schmalfeldt, 1992); and those that distinguish the categories according to whether they allow listeners to form expectations as to how they might end, called the *Prospective Schemas* model (Sears, 2015) (see §2). Experiment 2 directly compares these two models of cadential strength.

6.2.1 Analysis. To compare the mean IC estimates for the terminal events from each cadence category, each of the viewpoints was again analyzed for the terminal note events from the first violin and cello and the terminal chord events from the entire texture using a one-way ANOVA with a five-level between-groups factor called *cadence category* (PAC, IAC, HC, DC, and EV). To examine the potential differences in the IC estimates for the terminal events from each cadence category, each model includes two planned comparisons that do not assume equal variances, with a Bonferroni correction applied to the obtained statistics. In the first comparison, each level of *cadence category* was coded to represent two models of cadential strength: *Prospective Schemas* ($PAC \to IAC \to HC \to DC \to EV$) and *1-Schema* ($PAC \to IAC \to DC \to EV \to HC$). Polynomial contrasts with linear and quadratic terms were then included to estimate the goodness-of-fit for each model. In what follows, we report the contrast whose linear or quadratic trend accounts for the greatest proportion of variance in the outcome variable. The second comparison examines the hypothesis that the genuine cadence categories in Caplin's typology elicit lower IC estimates on average than the cadential deviations

(*Genuine* vs. *Deviations*).

6.2.2 Results. Figure 4 displays line plots of the mean IC estimates for the terminal note event in the first violin (left) and cello (right) for each level of *cadence category*. Table 6 presents the omnibus statistics and planned comparisons. For the first violin, the mean IC estimates revealed a main effect for the viewpoints `cpitch`, `csd`, and `selection`, but not for `melint`. Moreover, the best-fitting polynomial contrast revealed a positive (increasing) linear trend in the *Prospective Schemas* model (i.e., from the PAC to the EV categories) for every viewpoint model. The genuine cadence categories also received lower mean IC estimates than the cadential deviations in every model.

[Insert Figure 4 about here.]

For the cello, the IC estimates also revealed a main effect of *cadence category* for every viewpoint model, and the *Prospective Schemas* model again produced the best fit, with polynomial contrasts revealing positive quadratic trends for `cpitch` and `melint`, but positive linear trends for `csd` and `selection`. The quadratic trend exhibited in the `cpitch` and `melint` models for the cello probably reflects the statistical preference for smaller melodic intervals in the corpus, resulting in lower mean IC estimates for categories that feature stepwise motion in the bass (HC and DC), and higher estimates for categories featuring large leaps (PAC, IAC, and EV). This trend was not demonstrated in the `csd` and `selection` viewpoint models, however, as the DC category received higher IC estimates relative to the other categories in these models, thereby resulting in positive linear trends. Presumably, the HC category received the lowest IC estimates on average because scale-degree successions like $\hat{4}$–$\hat{5}$ are more common than successions like $\hat{5}$–$\hat{1}$. And yet successions like $\hat{5}$–$\hat{6}$ are also evidently *less* common than $\hat{5}$–$\hat{1}$, hence the higher IC estimates for the DC category and the increasing linear trend, $PAC \rightarrow IAC \rightarrow DC \rightarrow EV$. Finally, as expected, the genuine cadence categories received lower mean IC estimates than the cadential deviations in every model.

The left line plot in Figure 5 displays the mean IC estimates for the terminal chord event—represented by `vintcc` and `csdc`—for each level of *cadence category*. As before, the IC estimates revealed a main effect for `vintcc` and `csdc`, and the best-fitting polynomial contrast revealed a positive linear trend in the *Prospective Schemas* model for both models. The genuine cadence categories also received lower mean IC estimates than the cadential deviations for `vintcc` and `csdc`.

It is also noteworthy that the terminal events from the EV category generally received lower IC estimates than those from the DC category. Recall that evaded cadences are typically characterized not by a deviation in the harmonic progression (though such a deviation may take place), but rather by a sudden interruption in the projected resolution of the melody. In this collection, 10 of the 11 evaded cadences feature tonic harmony either in root position or in first inversion at the moment of cadential arrival. Given how often this harmony appears in the corpus, it is therefore not too surprising that the mean IC estimates decreased from the DC to the EV category.

The right line plot in Figure 5 displays the mean IC estimates for the terminal `composite` event for *cadence category*. In this case, the best-fitting polynomial contrast revealed a positive linear trend for the *Prospective Schemas* model. Thus, the model $PAC \rightarrow IAC \rightarrow HC \rightarrow DC \rightarrow EV$ accounted for roughly 55% of the variance in the mean IC estimates for `composite`, which represents the largest effect demonstrated across all of the polynomial contrasts from every viewpoint model. Finally, the genuine cadence categories again received lower mean IC estimates than the cadential deviations.

[Insert Figure 5 about here.]

6.2.3 Discussion. The mean IC estimates from IDyOM provide strong evidence in support of the *Prospective Schemas* model of cadential strength. Polynomial contrasts revealed significant positive linear trends for the viewpoints `vintcc`, `csdc`, `composite`, $\texttt{csd}_{\text{vc}}$, and $\texttt{selection}_{\text{vc}}$, as well as significant positive quadratic trends for $\texttt{cpitch}_{\text{vc}}$, $\texttt{melint}_{\text{vc}}$, and all of the viewpoints for the first violin. Furthermore, the claim that the genuine cadence categories elicit the strongest and most specific schematic expectations appears to be well supported by the experimental results from the second planned comparison, which revealed that the terminal events from the genuine cadence categories produced the lowest IC estimates on average for the viewpoint models from the first violin and across the entire texture, whereas the cadential deviations generally received the highest IC estimates on average.

Taken together, the reported findings support the role for expectancy in models of cadential strength, with the most complete or closed cadences also serving as the most expected. What is more, the results obtained here replicate the pattern of results reported by Sears, Caplin, and McAdams (2014). In that study, participants indicated how complete they found the end of each of a series of cadential excerpts from Mozart's keyboard sonatas. The genuine cadence categories and the cadential deviations received the highest and lowest completion ratings, respectively, which in light of the present findings suggests that the perceived strength of the cadential ending corresponds with the strength of the schematic expectations it generates in prospect. But recall that the perception of closure also depends on the cessation of expectations *following* the terminal events of the cadence. That is, the strength of the potential boundary between two sequential events results in part from the *increase* in information content (or decrease in probability) from the first to the second event (i.e., the last event of one group to the first event of the following group). The preceding analyses examined terminal events from cadential and non-cadential contexts in isolation, so Experiment 3 considers the role played by schematic expectations in boundary perception and event segmentation by examining the time course of IC estimates surrounding the terminal events of the cadence.

6.3 Experiment 3

Experiment 3 examines two claims about the relationship between expectancy and boundary perception: (1) that the terminal event of a group is the most expected (i.e., predictable) event in the surrounding sequence; and (2) that the next event in the sequence is comparatively unexpected (i.e., unpredictable). Again, the hypothesis here is that unexpected events engender prediction errors that lead the perceptual system to segment the event stream into discrete chunks (Kurby & Zacks, 2008). The terminal events from the PAC, IAC, and HC categories should be highly predictable, and prediction errors for the comparatively unpredictable events that follow should force listeners to segment the preceding cadential material. For the cadential deviations, however, prediction errors should occur at, rather than following, the terminal events of the cadence.

6.3.1 Analysis. In addition to the between-groups factor of *cadence category*, Experiment 3 includes a between-groups factor of *time* that consists of three levels: e_t, which represents the terminal event of the group, and e_{t-1} and e_{t+1}, which represent the immediately surrounding events. With more complex designs like this one, the number of significance tests can become prohibitively large, so we have restricted the investigation to the four viewpoints that serve as reasonable approximations of the two-voice framework characterizing the classical cadence: $\texttt{selection}_{\mathrm{vl1}}$ and $\texttt{selection}_{\mathrm{vc}}$ to represent the soprano and bass, respectively, and `vintcc` and `csdc` to each represent the entire texture. Experiment 3 analyzes these viewpoints using a 5×3 two-way ANOVA with between-groups factors of *cadence category* (PAC, IAC, HC, DC, and EV), and *time* (e_{t-1}, e_t, e_{t+1}).

By moving from one to two between-groups factors, the number of omnibus statistics and planned comparisons necessarily increases, and since Levene's test also revealed heteroscedastic groups for all four of the 5×3 viewpoint ANOVAs, the risk of committing a Type I error is considerably greater here than in the previous experiments. In this case, the two hypotheses mentioned above concern the interaction between *cadence category* and *time*: namely, whether the IC estimates for each cadence category increase or decrease significantly from one event to the next. Thus, the following analysis ignores the main effects and concentrates only on the interaction term of the two-way ANOVA. If the interaction is significant, we report simple main effects, which represent one-way ANOVAs with *time* as a factor for each level of cadence category. Finally, to examine the potential differences in the IC estimates for the levels of *time* for each cadence category, each simple main effect included two planned comparisons with Bonferroni correction that do not assume equal variances: (1) whether the IC estimate for e_t is lower on average than the surrounding events, e_{t-1} and e_{t+1} (e_t vs. *Surrounding*); and (2) whether the IC estimate for e_{t+1} is higher on average than the estimate for e_t (e_t vs. e_{t+1}).

6.3.2 Results. Figure 6 displays line plots of the mean IC estimates for the note events over time in the first violin (top) and cello (bottom) for each level of *cadence category*. Table 7 presents the omnibus statistics for the simple main effects and planned comparisons. To gain a more global picture of the IC time course, the line plots present the mean IC estimates for the seven-event sequence surrounding the terminal event of each cadence category, but the ANOVA models only consider the three events, -1, 0 (or Terminus), and 1. For the first violin, a two-way ANOVA of the mean IC estimates revealed a significant interaction between *cadence category* and *time*, $F(8, 718) = 3.88$, $p < .001$, est. $\omega^2 = .03$. The mean IC estimates for each level of *cadence category* revealed simple main effects for *PAC* and *HC*, but the remaining categories were not significant.

[Insert Figure 6 about here.]

Despite the non-significant simple main effects for the IAC, DC, and EV categories, simple planned comparisons revealed significant trends over time for every cadence category. As expected, the terminal event in the first violin received lower IC estimates on average than the immediately surrounding events for the genuine cadence categories and the DC category, though the latter trend was marginal. Thus, for cadences featuring melodies that resolve to presumably stable scale degrees like $\hat{1}$ or $\hat{3}$, the terminal event of the group is also the most predictable event in the sequence.

For the PAC, IAC, HC, and DC categories, the mean IC estimates increased significantly from e_t to e_{t+1}, thereby supporting the view that the strength of the perceptual boundary depends on the increase in information content following the terminal event of the cadence. And yet since the EV category replaces the expected terminal event in the melody with material that clearly initiates the subsequent process—often by leaping up to an unexpected scale degree like $\hat{5}$—one might therefore predict that a significant increase in information content should occur *at* (and not following) the expected terminal event of the group. This is exactly what we observe, with the expected terminal events from the EV category receiving the highest mean IC estimate in the sequence (see Table 7). Thus, the pattern of results from $\texttt{selection}_{\text{vl1}}$ is entirely consistent with the two main hypotheses: (1) the terminal event of a group is the most predictable event in the sequence, and (2) the next event is comparatively unpredictable. Here, the mean IC estimates for the first violin increased significantly following the predicted boundary for every cadence category in the collection.

For the cello, a two-way ANOVA of the mean IC estimates revealed a significant interaction between *cadence category* and *time*, $F(8, 717) = 13.02$, $p < .001$, est. $\omega^2 = .12$. Excepting *IAC*, the mean IC estimates also revealed simple main effects for every level of *cadence category*. As expected, the terminal event in the cello received lower IC estimates on average than the immediately surrounding events for the HC category, but the trend was reversed for the PAC, DC, and EV categories, and the trend for the IAC category was not significant.

For the HC category at least, the terminal event was also the most predictable event in the sequence. Furthermore, the significant increase in information content in the cello at the expected terminal event in the DC and EV categories is consistent with the behavior of cadential deviations. For the former category, the bass typically resolves deceptively to scale degrees like $\hat{6}$, thereby violating expectations for $\hat{1}$, whereas the latter category evades the expected resolution by leaping to other scale degrees to support harmonies like I^6. The significant increase in information content for the terminal event of the PAC category is somewhat more surprising, however. Recall from Experiment 2 that the mean IC estimates for the terminal events from each cadence category in the cello demonstrated a positive quadratic trend, with the HC category receiving the lowest IC estimates (see Figure 4). In that case, we suggested that small melodic intervals appear more abundantly in the corpus than large intervals, resulting in higher IC estimates for categories featuring large leaps (PAC, IAC, and EV). From this point of view, it seems reasonable that the mean IC estimates for the cello would increase at e_t for categories featuring large leaps or unexpected scale-degree continuations, as is the case with the PAC, IAC, DC, and EV categories.

Given this pattern of results for the cello, it should also come as little surprise that *HC* was the only category to demonstrate a significant increase in information content following the terminal event of the cadence. To be sure, the IC estimates for the cello did not significantly increase at e_{t+1} for the PAC and IAC categories, thereby undermining the hypothesis that for the genuine cadence categories at least, the perceived boundary follows the terminal events of the cadence. When the results from the first violin and the cello are considered together, *HC* was also the only category for which the IC estimates from $\texttt{selection}_{\text{vl1}}$ and $\texttt{selection}_{\text{vc}}$ decreased at e_t and then increased at e_{t+1}. If the PAC and IAC categories also generate strong and specific melodic *and* harmonic expectations for the terminal events of the cadence, the viewpoint models representing both voices of the two-voice framework should demonstrate congruent behavior.

Figure 7 displays line plots of the mean IC estimates for `vintcc` (top) and `csdc` (bottom) over time for each level of *cadence category*. Two-way ANOVAs of the mean IC estimates revealed a significant interaction between *cadence category* and *time* for both viewpoint models (`vintcc`, $F(8, 719) = 3.13$, $p = .002$, est. $\omega^2 = .03$; `csdc`, $F(8, 704) = 2.99$, $p = .003$, est. $\omega^2 = .03$). Simple main effects and planned comparisons were not significant for `csdc`, however, so it will not be reported here. For `vintcc`, the mean IC estimates revealed simple main effects for the genuine cadence categories, but not for the cadential deviations. As expected, the terminal chord event received lower IC estimates on average compared to the surrounding events for the genuine cadence categories. Although the trend was reversed for the cadential deviations, with the mean IC estimates increasing from e_{t-1} to e_t, the difference was not significant for either category. Finally, the mean IC estimates increased significantly from e_t to e_{t+1} for *PAC* and *HC*, but this trend was marginal for *IAC*.

[Insert Figure 7 about here.]

6.3.3 Discussion. The viewpoint model for `vintcc` demonstrated a similar trend to that found in $\texttt{selection}_{\text{vl1}}$ for the genuine cadence categories, with the mean IC estimates decreasing from e_{t-1} to e_t, and then increasing from e_t to e_{t+1}. These two viewpoint models also displayed congruent behavior for the EV category, with both models increasing from e_{t-1} to e_t, suggesting that the perceptual boundary precedes (rather than follows) the expected terminal event in evaded cadences. For the DC category, parametric noncongruence obtained, with the mean IC estimates at e_t decreasing in $\texttt{selection}_{\text{vl1}}$ but increasing in `vintcc`. Thus, across the levels of *cadence category* and *time*, the $\texttt{selection}_{\text{vl1}}$ and `vintcc` viewpoint models supported our initial hypotheses: (1) that the terminal event of a group is the most expected (i.e., predictable) event in the sequence; and (2) that the next event is comparatively unexpected (i.e., unpredictable).

7 Conclusions

This study examined three claims about the relationship between expectancy and cadential closure: (1) terminal events from cadential contexts are more predictable than those from non-cadential contexts; (2) models of cadential strength advanced in cadence typologies like the one employed here reflect the formation, violation, and fulfillment of schematic expectations; and (3) a significant decrease in predictability follows the terminal note and chord events of the cadential process. To that end, we created a corpus of Haydn string quartets to serve as a proxy for the musical experiences of listeners situated in the classical style, selected a number of viewpoints to represent suitable (i.e., cognitively plausible) representations of the musical surface, and then employed IDyOM—an *n*-gram model that predicts the next note or chord event in a musical stimulus through unsupervised learning of sequential structure—to simulate the formation of schematic expectations during music listening.

The findings from Experiment 1 indicate that the terminal note and chord events from perfect authentic cadences are more predictable than (1) non-cadential events featuring tonic harmony in root position and supporting any scale degree in the soprano, and (2) non-cadential events featuring any other harmony and any other scale degree in the soprano. For half cadences, significant effects were limited to the chord models (`vintcc`, `csdc`, and `composite`) and the `csd` viewpoint model, but the terminal events from half-cadential contexts were still more predictable than those from non-cadential contexts. Experiment 2 provided strong evidence in support of the *Prospective Schemas* model of cadential strength (PAC→IAC→HC→DC→EV), with the genuine cadence categories (PAC, IAC, HC) and cadential deviations (DC, EV) in Caplin's typology eliciting the lowest and highest IC estimates on average, respectively. Finally, the results from Experiment 3 indicated that unexpected events—like those directly following the terminal note and chord events from genuine cadences—engender prediction errors that presumably lead the perceptual system to segment the event stream immediately following the cadential process.

Taken together, the reported findings support the role of expectancy in models of cadential closure, with the most complete or closed cadences also serving as the most expected or probable. Nevertheless, future studies will need to address a number of limitations in the current investigation. First, the rather meager sample size for three of the five cadence categories in the collection—as well as the corpus more generally—casts some doubt upon the generalizability of the reported findings. That the estimates from IDyOM correspond so well with theoretical predictions suggests that these findings may be robust to issues of sample size, but future studies should look to expand the collection considerably, as well as to consider how the relationship between expectancy and cadential closure varies for other genres and style periods.

Second, we selected individual viewpoints if existing theoretical or experimental evidence justified their inclusion, such as `melint` and `csd` in melodic contexts (Dowling, 1981; Krumhansl, 1990), and `vintcc` and `csdc` in harmonic contexts (Gjerdingen, 2007). Yet in a few instances, `cpitch`—which was only included among the melodic viewpoints to serve as a baseline for model comparison—produced similar results (see Experiment 2). One reason for this finding is that many of the viewpoints characterizing melodic organization in Haydn's string quartets systematically covary such that statistical regularities governing the more cognitively plausible viewpoints (e.g., `melint` and `csd`) also appear in the less plausible ones (`cpitch`), albeit more weakly. In a melody composed in the key of C-major, for example, B♮ presumably functions as the *leading tone*, not because the twelve-tone chromatic universe regularly features this two-note sequence regardless of the particular tonal context, but because C♮ typically follows B♮ in the key of C-major, forming one of the many statistical associations characterizing the tonal system. Nevertheless, the tendency for small melodic intervals and the prevalence of certain keys in tonal music—wherein B♮ is more likely to progress to C♮ than to, say, A♭—ensures that IDyOM will learn statistical regularities in basic viewpoints like `cpitch` that are correlated to those found in other melodic viewpoints like `melint` or `csd`.

What is more, rather than assume that listeners expect *specific* intervals in a melodic sequence, as is IDyOM's approach using `melint`, existing models of melodic expectation typically theorize that listeners expect smaller melodic intervals regardless of the preceding context, a principle known as *pitch proximity* (e.g., Cuddy & Lunney, 1995; Margulis, 2005; Narmour, 1990; Schellenberg, 1997). Yet in this study, we only assume that listeners form expectations on the basis of statistical regularities among melodic intervals in a sequence. Vos and Troost (1989) have demonstrated, for example, that small intervals are far more common than large intervals in Western tonal music, so it should not be surprising that IDyOM produces higher probability estimates for smaller intervals, just as do proximity-based models. The difference in these two approaches is thus theoretical, rather than empirical, in that IDyOM bases its predictions on a theory of implicit statistical learning, whereas proximity-based models also sometimes base their assumptions on other (sensory or psychoacoustic) mechanisms. It is certainly possible that these mechanisms influence the preference for smaller over larger intervals in Western tonal music—or indeed, the formation of expectations during music listening more generally—but we do not examine this assumption here.

Perhaps more importantly, the cross entropy estimates for the melodic models in this study indicate that IDyOM was more certain about its predictions for more cognitively plausible viewpoints like `melint` and `csd` than for less plausible ones like `cpitch`, thereby reinforcing the view that representations of the musical surface are 'cognitively plausible' to the degree that they minimize prediction errors for future events (Pearce & Wiggins, 2012). Indeed, if prediction is the 'primary function' of the brain (Hawkins & Blakeslee, 2004, p. 89), and listeners learn to compress information during processing by only retaining representations of the musical surface that minimize uncertainty (Pearce & Wiggins, 2012), it therefore seems reasonable to include `cpitch` and `melint` among a potentially large number of candidate viewpoints in the initial model configuration and allow IDyOM to select the viewpoint (or combination of viewpoints) that minimizes uncertainty empirically (i.e., in an unsupervised manner). Furthermore, in this case IDyOM benefited from human annotations of tonal information in `csd`, but future studies could employ viewpoints like the General Chord Type (GCT) representation (Cambouropoulos, 2015), which automatically produces Roman numeral-like encodings of complex polyphonic corpora.

Third, IDyOM's modeling architecture could be further improved to more closely resemble the mechanisms by which listeners form expectations for future events. A number of studies in the language modelling literature have demonstrated the utility of non-contiguous n-grams for the discovery and classification of recurrent patterns (i.e., collocations) (Guthrie, Allison, Liu, Guthrie, & Wilks, 2006; Huang, Beeferman, & Huang, 1999; Simons, Ney, & Martin, 1997), but the present investigation was limited to contiguous *n*-grams. Creel, Newport, and Aslin (2004) have shown, for example, that listeners can learn non-contiguous statistical regularities in melodic sequences if the intervening events are segregated in terms of pitch height or timbre. What is more, Gjerdingen (2014) has suggested that for stimuli demonstrating hierarchical structure, non-contiguous events often serve as focal points in the syntax. This problem is particularly acute for corpus studies of tonal harmony, where the musical surface contains considerable repetition, and many of the vertical sonorities from the notated score do not represent triads or seventh chords, thereby obscuring the most recurrent patterns. IDyOM is presently capable of including non-contiguous *n*-grams using *threaded* viewpoints, which sample events from a *base* viewpoint like `cpitch` according to some *test* viewpoint that represents positions in the sequence, such as metric downbeats or phrase boundaries. Pearce (2005) has shown that these viewpoints improve model performance in melodic prediction tasks, so it is possible that they may also improve model predictions for the terminal events from cadential contexts.

Finally, the present approach depended entirely on simulation. If the brain is a 'prediction machine' that generates expectations about future events by forming associations between co-occurring attributes within the external environment, as some have suggested (Bar, 2007; Clark, 2013), then behavioral and neural manifestations of expectancy formation, violation, and fulfillment should correspond in some way with the model simulations reported in this study. In this case, the model estimates generated by IDyOM support the view that cadences and other recurrent closing patterns serve as the most predictable, probabilistic, specifically envisaged formulæ in all of tonal music (Huron, 2006; Meyer, 1956). To demonstrate further that the schematic expectations formed by listeners for cadences and other recurrent temporal patterns amount to these sorts of probabilistic inferences requires an entirely different approach, one in which the listener, rather than the music, represents the primary object of study.

References

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* London, United Kingdom: Cambridge University Press.

Blombach, A. (1987). Phrase and cadence: A study of terminology and definition. *Journal of Music Theory Pedagogy*, *1*, 225–251.

Brent, M. R. (1999). An efficient, probabilistically sound algorithm for segmentation and word discovery. *Machine Learning*, *34*, 71–105.

Brown, P. F., Della Pietra, V. J., Della Pietra, S. A., Lai, J. C., & Mercer, R. L. (1992). An estimate of an upper bound for the entropy of English. *Computational Linguistics*, *18*(1), 31–40.

Bunton, S. (1996). *On-line stochastic processes in data compression* (Unpublished doctoral dissertation). University of Washington, Seattle, WA.

Bunton, S. (1997). Semantically motivated improvements for PPM variants. *The Computer Journal*, *40*(2/3), 76–93.

Cambouropoulos, E. (2015). Computational music analysis. In D. Meredith (Ed.), (pp. 31–56). Heidelberg: Springer International Publishing.

Caplin, W. E. (1998). *Classical form: A theory of formal functions for the instrumental music of Haydn, Mozart, and Beethoven.* New York: Oxford University Press.

Caplin, W. E. (2004). The classical cadence: Conceptions and misconceptions. *Journal of the American Musicological Society*, *57*(1), 51–118.

Chen, S. F., & Goodman, J. (1999). An empirical study of smoothing techniques for language modeling. *Computer Speech & Language*, *13*, 359–394.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 1–73.

Cleary, J. G., & Teahan, W. J. (1997). Unbounded length contexts for PPM. *The*

Computer Journal, *40*(2/3), 67–75.

Cleary, J. G., & Witten, I. H. (1984). Data compression using adaptive coding and partial string matching. *IEEE Transactions on Communications*, *32*(4), 396–402.

Cohen, B. H. (2008). *Explaining psychological statistics*. Hoboken, NJ: John Wiley & Sons, Inc.

Cohen, P., Adams, N., & Heeringa, B. (2007). Voting experts: An unsupervised algorithm for segmenting sequences. *Intelligent Data Analysis*, *11*, 607–625.

Conklin, D. (1988). Modelling and generating music using multiple viewpoints. In *Proceedings of the first workshop on artificial intelligence and music* (pp. 125–137). St. Paul, MN: The American Association for Artificial Intelligence.

Conklin, D. (1990). *Prediction and entropy of music* (Unpublished master's thesis). University of Calgary, Calgary, AB.

Conklin, D. (2002). Representation and discovery of vertical patterns in music. In C. Anagnostopoulou, M. Ferrand, & A. Smaill (Eds.), *Music and artifical intelligence: Proc. icmai 2002* (Vol. 2445, pp. 32–42). Springer-Verlag.

Conklin, D., & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, *24*(1), 51–73.

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(5), 1119–1130.

Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception and Psychophysics*, *57*(4), 451–462.

Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, *10*(7), 1895–1923.

Dowling, W. J. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, *1*, 30–49.

Dunsby, J. (1980). Schoenberg on cadence. *Journal of the Arnold Schoenberg Institute*, *4*(1), 41–49.

Eberlein, R. (1997). A method of analysing harmony, based on interval patterns or "Gestalten". In M. Leman (Ed.), *Music, gestalt, and computing* (Vol. 1317, pp. 225–236). Springer Berlin / Heidelberg.

Eberlein, R., & Fricke, J. (1992). *Kadenzwahrnehmung und kadenzgeschichte: Ein beitrag zu einer grammatik der musik*. Frankfurt/M.: P. Lang.

Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective, and Behavioral Neuroscience*, *13*(2).

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Gjerdingen, R. O. (1988). *A classic turn of phrase: Music and the psychology of convention*. Philadelphia: University of Pennsylvania Press.

Gjerdingen, R. O. (1991). Defining a prototypical utterance. *Psychomusicology*, *10*(2), 127–139.

Gjerdingen, R. O. (2007). *Music in the galant style*. New York: Oxford University Press.

Gjerdingen, R. O. (2014). "Historically informed" corpus studies. *Music Perception*, *31*(3), 192–204.

Guthrie, D., Allison, B., Liu, W., Guthrie, L., & Wilks, Y. (2006). A closer look at skip-gram modelling. In *Proceedings of the 5th international conference on language resources and evaluation (lrec 06)* (p. 1222-1225). European Language Resources Association.

Hansen, N. C., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology*, *5*, 1–17. doi: 10.3389/fpsyg.2014.1052

Hawkins, J., & Blakeslee, S. (2004). *On intelligence*. New York, NY: Times Books.

Hedges, G. A., Thomas Wiggins. (2016). The prediction of merged attributes with multiple viewpoint systems. *Journal of New Music Research*. doi:

10.1080/09298215.2016.1205632

Hepokoski, J., & Darcy, W. (2006). *Elements of sonata theory: Norms, types, and deformations in the late-eighteenth-century sonata.* New York: Oxford University Press.

Huang, M., Beeferman, D., & Huang, X. D. (1999). Improved topic-dependent language modeling using information retrieval techniques. In *Proceedings of the international conference on acoustics, speech and signal processing.* Washington, DC: IEEE Computer Society Press.

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

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

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–79.

Latham, E. (2009). Drei nebensonnen: Forte's linear-motivic analysis, Korngold's *Die Tote Stadt*, and Schubert's *Winterreise* as visions of closure. *Gamut*, *2*(1), 299–345.

MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms.* Cambridge, UK: Cambridge University Press.

Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing.* Cambridge, MA: MIT Press.

Margulis, E. H. (2003). *Melodic expectation: A discussion and model* (Unpublished doctoral dissertation). Columbia University, New York, NY.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception*, *22*(4), 663–714.

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

Meyer, L. B. (1967). *Music, the arts, and ideas.* Chicago: The University of Chicago Press.

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

Moffat, A. (1990). Implementing the PPM data compression scheme. *IEEE Transactions on Communications*, *COM-38*(11), 1917–1921.

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

Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. *Neuropsychologia*, *50*, 1483–1493.

Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia*, *51*, 1749–1762.

Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition* (Unpublished doctoral dissertation). City University, London.

Pearce, M. T., Conklin, D., & Wiggins, G. A. (2005). Methods for combining statistical models of music. In U. K. Wiil (Ed.), *Computer music modelling and retrieval* (pp. 295–312). Heidelberg, Germany: Springer Verlag.

Pearce, M. T., Müllensiefen, D., & Wiggins, G. A. (2010). The role of expectation and probabilistic learning in auditory boundary perception: A model comparison. *Perception*, *39*, 1367–1391.

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, *50*, 302–313.

Pearce, M. T., & Wiggins, G. A. (2004). Improved methods for statistical modelling of monophonic music. *Journal of New Music Research*, *33*(4), 367–385.

Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception*, *23*(5), 377–405.

Pearce, M. T., & Wiggins, G. A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, *4*, 625–652.

Peebles, C. (2011). *The role of segmentation and expectation in the perception of closure* (Unpublished doctoral dissertation). Florida State University, Tallahassee, FL.

Piston, W. (1962). *Harmony* (3rd ed. ed.). New York: W. W. Norton & Company.

Potter, K., Wiggins, G. A., & Pearce, M. T. (2007). Towards greater objectivity in music theory: Information-dynamic analysis of minimalist music. *Musicae Scientiae*, *11*(2), 295–324.

Quinn, I. (2010). Are pitch-class profiles really key for key. *Zeitschrift der Gesellschaft der Musiktheorie*, *7*, 151–163.

Ratner, L. G. (1980). *Classic music: Expression, form, and style.* New York: Schirmer Books.

Rohrmeier, M., & Neuwirth, M. (2011, January). Characteristic signatures of cadences in Mozart's piano sonatas. In M. Neuwirth and P. Bergé (Chairs), *What is a cadence? Theoretical and analytical perspectives on cadences in the classical repertoire.* Symposium conducted at the Academia Belgica, Rome.

Rosen, C. (1972). *The classical style: Haydn, mozart, beethoven.* New York: Norton.

Rosner, B., & Narmour, E. (1992). Harmonic closure: Music theory and perception. *Music Perception*, *9*(4), 383–412.

Schellenberg, E. G. (1997). Simplifying the implication-realization model of melodic expectancy. *Music Perception*, *14*(3), 295–318.

Schmalfeldt, J. (1992). Cadential processes: The evaded cadence and the 'one more time' technique. *Journal of Musicological Research*, *12*(1), 1–52.

Sears, D. (2015). The perception of cadential closure. In M. Neuwirth & P. Bergé (Eds.), *What is a cadence? theoretical and analytical perspectives on cadences in the classical repertoire* (p. 251-283). Leuven: Leuven University Press.

Sears, D. (2016). *The classical cadence as a closing schema: Learning, memory, and* *perception* (Unpublished doctoral dissertation). McGill University, Montreal, Canada.

Sears, D., Caplin, W. E., & McAdams, S. (2014). Perceiving the classical cadence. *Music Perception*, *31*(5), 397-417.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379-423.

Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell System Technical Journal*, *30*, 50–64.

Simons, M., Ney, H., & Martin, S. C. (1997). Distant bigram language modeling using maximum entropy. In *Proceedings of the international conference on acoustics, speech and signal processing* (pp. 787–790). Washington, DC: IEEE Computer Society Press.

Snyder, B. (2000). *Music and memory: An introduction.* Cambridge, MA: The MIT Press.

Temperley, D. (2004). *The cognition of basic musical structures.* Cambridge, MA: The MIT Press.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception*, *6*(4), 383-396.

Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, *38*(3/4), 330–336.

Witten, I. H., & Bell, T. C. (1991). The zero-frequency problem: Estimating the probabilities of novel events in adaptive text compression. *IEEE Transactions on Information Theory*, *37*(4), 1085–1094.

Footnotes

[1]For basic types like `cpitch`, Ψ_τ is simply a projection function, thereby returning as output the same values it receives as input (Pearce, 2005, p. 59).

[2]Context models like the one just described also often use a technique called *exclusion*, which improves the final probability estimate by reclaiming a portion of the probability mass in lower-order models that is otherwise wasted on redundant predictions (i.e., the counts for events that were predicted in the higher-order distributions do not need to be included in the calculation of the lower-order distributions).

[3]For a worked example of the PPM* method, see Sears (2016).

[4]`http://kern.ccarh.org/`.

Table 1

The cadential types and categories, along with the harmonic and melodic characteristics and the count for each category in the cadence collection. Categories marked with an asterisk are cadential deviations.

Types	*Categories*	*Essential Characteristics*	*N*
I	Perfect Authentic (PAC)	V – I $\hat{1}$	122
	Imperfect Authentic (IAC)	V – I $\hat{3}$ or $\hat{5}$	9
	Deceptive (DC)*	V – ?, Typically vi $\hat{1}$ or $\hat{3}$	19
	Evaded (EV)*	V – ? ?, Typically $\hat{5}$	11
V	Half (HC)	? – V $\hat{5}$, $\hat{7}$, or $\hat{2}$	84

Table 2

Reference information (Opus number, work, movement, measures), keys (case denotes mode), time signatures, and tempo markings for the exposition sections in the corpus.

Excerpt	*Key*	*Time Signature*	*Tempo Marking*
Op. 17, No. 1, i, mm. 1–43	E	4/4	Moderato
Op. 17, No. 2, i, mm. 1–38	F	4/4	Moderato
Op. 17, No. 3, iv, mm. 1–26	E♭	4/4	Allegro molto
Op. 17, No. 4, i, mm. 1–53	c	4/4	Moderato
Op. 17, No. 5, i, mm. 1–33	G	4/4	Moderato
Op. 17, No. 6, i, mm. 1–73	D	6/8	Presto
Op. 20, No. 1, iv, mm. 1–55	E♭	2/4	Presto
Op. 20, No. 3, i, mm. 1–94	g	2/4	Allegro con spirito
Op. 20, No. 3, iii, mm. 1–43	G	3/4	Poco Adagio
Op. 20, No. 3, iv, mm. 1–42	g	4/4	Allegro molto
Op. 20, No. 4, i, mm. 1–112	D	3/4	Allegro di molto
Op. 20, No. 4, iv, mm. 1–49	D	4/4	Presto scherzando
Op. 20, No. 5, i, mm. 1–48	f	4/4	Allegro moderato
Op. 20, No. 6, ii, mm. 1–27	E	cut	Adagio
Op. 33, No. 1, i, mm. 1–37	b	4/4	Allegro moderato
Op. 33, No. 1, iii, mm. 1–40	D	6/8	Andante
Op. 33, No. 2, i, mm. 1–32	E♭	4/4	Allegro moderato
Op. 33, No. 3, iii, mm. 1–29	F	3/4	Adagio
Op. 33, No. 4, i, mm. 1–31	B♭	4/4	Allegro moderato
Op. 33, No. 5, i, mm. 1–95	G	2/4	Vivace assai
Op. 33, No. 5, ii, mm. 1–30	g	4/4	Largo
Op. 50, No. 1, i, mm. 1–60	B♭	cut	Allegro
Op. 50, No. 1, iv, mm. 1–75	B♭	2/4	Vivace
Op. 50, No. 2, i, mm. 1–106	C	3/4	Vivace
Op. 50, No. 2, iv, mm. 1–86	C	2/4	Vivace assai
Op. 50, No. 3, iv, mm. 1–74	E♭	2/4	Presto
Op. 50, No. 4, i, mm. 1–64	f♯	3/4	Allegro spirituoso
Op. 50, No. 5, i, mm. 1–65	F	2/4	Allegro moderato
Op. 50, No. 5, iv, mm. 1–54	F	6/8	Vivace
Op. 50, No. 6, i, mm. 1–54	D	4/4	Allegro
Op. 50, No. 6, ii, mm. 1–25	d	6/8	Poco Adagio
Op. 54, No. 1, i, mm. 1–47	G	4/4	Allegro con brio
Op. 54, No. 1, ii, mm. 1–54	C	6/8	Allegretto
Op. 54, No. 2, i, mm. 1–87	C	4/4	Vivace
Op. 54, No. 3, i, mm. 1–58	E	cut	Allegro
Op. 54, No. 3, iv, mm. 1–82	E	2/4	Presto
Op. 55, No. 1, ii, mm. 1–36	D	2/4	Adagio cantabile
Op. 55, No. 2, ii, mm. 1–76	f	cut	Allegro
Op. 55, No. 3, i, mm. 1–75	B♭	3/4	Vivace assai
Op. 64, No. 3, i, mm. 1–69	B♭	3/4	Vivace assai
Op. 64, No. 3, iv, mm. 1–79	B♭	2/4	Allegro con spirito
Op. 64, No. 4, i, mm. 1–38	G	4/4	Allegro con brio
Op. 64, No. 4, iv, mm. 1–66	G	6/8	Presto
Op. 64, No. 6, i, mm. 1–45	E♭	4/4	Allegretto
Op. 71, No. 1, i, mm. 1–69	B♭	4/4	Allegro
Op. 74, No. 1, i, mm. 1–54	C	4/4	Allegro moderato
Op. 74, No. 1, ii, mm. 1–57	G	3/8	Andantino grazioso
Op. 76, No. 2, i, mm. 1–56	d	4/4	Allegro
Op. 76, No. 4, i, mm. 1–68	B♭	4/4	Allegro con spirito
Op. 76, No. 5, ii, mm. 1–33	F♯	cut	Largo. Cantabile e mesto

Table 3
Descriptive statistics for the corpus.

Instrumental Part	*N*	*M (SD)*	*Range*
Note Events			
Violin 1	14,506	290 (78)	133–442
Violin 2	10,653	213 (70)	69–409
Viola	9156	183 (63)	79–381
Cello	8463	169 (60)	64–326
Chord Events			
Expansion[a]	20,290	406 (100)	189–620

[a] To identify chord events in polyphonic textures, full expansion duplicates overlapping note events at every unique onset time (Conklin, 2002).

Table 4
Analysis of variance and planned comparisons predicting the information content estimates from all viewpoint models with tonic closure.

	Omnibus				**Comparisons**							
					PAC vs. *Non-Cadence*				*PAC* vs. *Tonic*			
Viewpoint	*df*	Welch F	est. ω^2	*p*	*df*	*t*	*r*	*p*	*df*	*t*	*r*	*p*
Note Events												
Violin 1												
`cpitch`	241.99	1.56	.003	*NS*	241.58	−1.43	.09	*NS*	241.99	−0.73	.05	*NS*
`melint`	238.65	5.80	.03	.003	289.44	−3.38	.19	.002	239.31	−2.47	.16	.029
`csd`	238.43	7.83	.04	<.001	292.69	−3.19	.18	.003	238.25	−1.23	.08	.220
`selection`	237.80	8.11	.04	<.001	297.25	−3.64	.21	<.001	237.52	−1.91	.12	*NS*
Cello												
`cpitch`	239.77	29.42	.13	<.001	286.12	7.46	.40	<.001	236.13	7.21	.42	<.001
`melint`	229.78	41.29	.18	<.001	342.08	9.04	.44	<.001	208.37	7.81	.48	<.001
`csd`	233.81	17.76	.08	<.001	319.18	4.31	.23	<.001	231.65	5.95	.36	<.001
`selection`	231.83	35.32	.16	<.001	333.12	7.90	.40	<.001	218.98	8.01	.48	<.001
Chord Events												
`vintcc`	232.01	26.32	.12	<.001	311.30	−5.53	.30	<.001	238.84	−0.96	.06	*NS*
`csd`	237.59	9.96	.05	<.001	296.92	−3.81	.22	<.001	238.39	−1.67	.11	*NS*
`composite`	238.02	13.61	.06	<.001	281.84	−4.59	.26	<.001	241.89	−2.24	.14	.052

Note. *NS* = non-significant. Planned comparisons corrected with Bonferroni adjustment.

Table 5

Analysis of variance and planned comparisons predicting the information content estimates from all viewpoint models with dominant closure.

	Omnibus				Comparisons							
					HC vs. *Non-Cadence*				*HC* vs. *Dominant*			
Viewpoint	*df*	Welch F	est. ω^2	*p*	*df*	*t*	*r*	*p*	*df*	*t*	*r*	*p*
Note Events												
Violin 1												
`cpitch`	164.25	4.82	.02	.009	191.26	−3.11	.22	.004	165.44	−2.55	.19	.024
`melint`	159.52	5.89	.03	.003	227.92	−3.42	.22	.002	146.00	−2.99	.24	.007
`csd`	162.09	8.14	.04	<.001	202.06	−3.80	.26	<.001	165.50	−2.16	.17	*NS*
`selection`	162.39	9.38	.04	<.001	209.13	−4.26	.28	<.001	161.61	−2.91	.22	.008
Cello												
`cpitch`	162.93	0.78	≈ 0	*NS*	209.18	1.24	.08	*NS*	157.94	0.94	.06	*NS*
`melint`	164.37	1.15	≈ 0	*NS*	195.71	1.46	.09	*NS*	163.31	0.97	.06	*NS*
`csd`	161.11	6.19	.03	.003	216.02	−3.36	.22	.002	159.74	−1.97	.25	*NS*
`selection`	162.88	1.37	.002	*NS*	207.10	−1.59	.10	*NS*	161.31	2.45	.16	*NS*
Chord Events												
`vintcc`	160.56	16.66	.08	<.001	216.71	−1.42	.09	*NS*	160.57	2.45	.16	*NS*
`csd`	164.84	4.00	.02	.020	187.47	−1.86	.12	*NS*	165.37	−0.36	.02	*NS*
`composite`	162.93	8.04	.04	<.001	203.00	−3.56	.24	<.001	163.99	−1.74	.13	*NS*

Note. *NS* = non-significant. Planned comparisons corrected with Bonferroni adjustment.

Table 6

Analysis of variance and planned comparisons predicting the information content estimates from all viewpoint models with cadence categories.

	Omnibus				**Comparisons**								
					Prospective Schemas					*Genuine* vs. *Deviations*			
Viewpoint	*df*	Welch *F*	est. ω^2	*p*	*Trend*	*df*	*t*	*r*	*p*	*df*	*t*	*r*	*p*
Note Events													
Violin 1													
`cpitch`	32.42	3.19	.03	.026	Linear	13.47	3.40	.68	.013	17.19	−3.64	.66	.006
`melint`	31.57	2.34	.02	*NS*	Linear	12.13	3.00	.65	.032	14.14	−2.94	.62	.033
`csd`	29.95	3.02	.03	.033	Linear	18.99	3.43	.62	.008	30.51	−3.17	.50	.009
`selection`	29.80	3.77	.04	.013	Linear	16.38	3.85	.69	.004	26.03	−3.66	.58	.003
Cello													
`cpitch`	31.18	8.86	.11	<.001	Quadratic	25.39	4.38	.66	<.001	27.62	−2.59	.44	.045
`melint`	30.72	6.81	.09	<.001	Quadratic	14.04	4.11	.74	.003	15.93	−2.93	.59	.030
`csd`	31.25	13.99	.18	<.001	Linear	14.83	4.09	.73	.003	24.39	−5.18	.72	.003
`selection`	30.66	14.99	.19	<.001	Linear	13.41	3.83	.72	.003	19.91	−4.95	.74	.003
Chord Events													
`vintcc`	29.94	6.68	.08	.001	Linear	16.96	3.56	.65	.007	28.01	−3.92	.60	.001
`csd`	31.90	7.02	.09	<.001	Linear	27.04	4.14	.62	<.001	39.39	−3.84	.52	.001
`composite`	30.91	7.21	.09	<.001	Linear	19.23	4.81	.74	<.001	31.87	−4.58	.63	.003

Note. *NS* = non-significant. Planned comparisons corrected with Bonferroni adjustment.

Table 7

Analysis of variance and planned comparisons predicting the information content estimates from **`selection`**$_{vl1}$*,* **`selection`**$_{vc}$*, and* **`vintcc`** *over time for each cadence category.*

	Omnibus				Comparisons							
					e_t vs. *Surrounding*				e_t vs. e_{t+1}			
Viewpoint	*df*	Welch *F*	est. ω^2	*p*	*df*	*t*	*r*	*p*	*df*	*t*	*r*	*p*
Note Events												
`selection`$_{vl1}$												
PAC	236.76	67.63	.35	<.001	298.45	−9.88	.50	<.001	223.50	11.64	.61	<.001
IAC	14.95	3.89	.02	*NS*	18.48	−2.79	.54	.024	12.53	2.78	.54	.032
HC	162.48	32.34	.20	<.001	209.41	−6.92	.43	<.001	160.95	8.02	.53	<.001
DC	35.86	2.84	.01	*NS*	36.92	−2.10	.33	.085	35.62	2.41	.55	.043
EV	18.67	5.22	.03	*NS*	14.87	2.54	.55	.043	19.23	−1.58	.34	*NS*
`selection`$_{vc}$												
PAC	227.67	42.14	.25	<.001	322.14	3.18	.17	.003	180.58	1.65	.12	*NS*
IAC	14.20	3.38	.02	*NS*	21.98	1.98	.39	*NS*	10.57	−0.71	.21	*NS*
HC	161.48	15.57	.11	<.001	213.66	−5.44	.35	<.001	151.11	3.76	.29	<.001
DC	35.50	8.93	.06	.004	43.02	2.98	.41	.010	33.15	−0.97	.17	*NS*
EV	19.87	8.29	.06	.012	17.49	4.06	.70	.002	19.22	−3.19	.59	.009
Chord Events												
`vintcc`												
PAC	235.55	9.74	.07	<.001	313.01	−3.63	.20	.001	220.21	4.42	.29	<.001
IAC	15.21	6.48	.04	.046	22.58	−3.49	.59	.004	13.71	2.09	.49	*NS*
HC	161.39	8.26	.06	.002	216.80	−4.07	.27	<.001	157.62	3.25	.25	.003
DC	34.73	2.05	.01	*NS*	38.02	1.14	.18	*NS*	34.51	0.05	.01	*NS*
EV	19.91	0.18	≈ 0	*NS*	18.50	0.25	.06	*NS*	19.98	0.04	.01	*NS*

Note. NS = non-significant. Planned comparisons corrected with Bonferroni adjustment.

Figure Captions

Figure 1. Top: First violin part from Haydn's String Quartet in E, Op. 17/1, i, mm. 1–2. Bottom: Viewpoint representation.

Figure 2. Top: Bar plots of the mean information content (IC) estimated for the terminal note event in the first violin (left) and cello (right) for each level of *tonic closure*. Viewpoints include `cpitch`, `melint`, `csd`, and an optimized combination called `selection`, which represents `melint` and the linked viewpoint $\texttt{csd} \otimes \texttt{cpitch}$. Bottom left: Bar plot of the mean information content (IC) estimated for the terminal `vintcc` and `csdc` for each level of *tonic closure*. Bottom right: Line plot of the mean information content (IC) estimated for the combination called `composite`, which represents the dot product of $\texttt{selection}_{\text{vl1}}$ and `csdc`. Whiskers represent ± 1 standard error.

Figure 3. Top: Bar plots of the mean information content (IC) estimated for the terminal note event in the first violin (left) and cello (right) for each level of *dominant closure*. Viewpoints include `cpitch`, `melint`, `csd`, and an optimized combination called `selection`, which represents `melint` and the linked viewpoint $\texttt{csd} \otimes \texttt{cpitch}$. Bottom left: Bar plot of the mean information content (IC) estimated for the terminal `vintcc` and `csdc` for each level of *dominant closure*. Bottom right: Line plot of the mean information content (IC) estimated for the combination called `composite`, which represents the dot product of $\texttt{selection}_{\text{vl1}}$ and `csdc`. Whiskers represent ± 1 standard error.

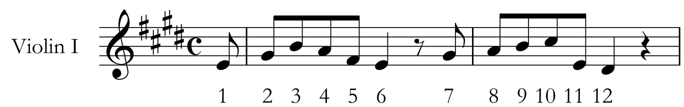
Figure 4. Line plots of the mean IC estimated for the terminal note event in the first violin (left) and cello (right) for each cadence category. Viewpoints include `cpitch`, `melint`, `csd`, and an optimized combination called `selection`, which represents `melint` and the linked viewpoint `csd`⊗`cpitch`. Whiskers represent ±1 standard error.

Figure 5. Left: Line plot of the mean IC estimated for the resolving chord event for each cadence category. Right: Line plot of the mean IC estimated for the combination called `composite`, which represents the product of $\texttt{selection}_{\text{vl1}}$ and `csdc`. Whiskers represent ±1 standard error.

Figure 6. Time course of the mean IC estimated for the events surrounding the terminal note event in the first violin (top) and cello (bottom) for each cadence category using the viewpoint `selection`, which represents `melint` and the linked viewpoint `csd` ⊗ `cpitch`. The statistical analysis pertains to event numbers −1, 0 (or Terminus), and 1. Whiskers represent ±1 standard error.

Figure 7. Time course of the mean IC estimated for the events surrounding the terminal chord event for `vintcc` (top) and `csdc` (bottom) for each cadence category. The statistical analysis pertains to event numbers −1, 0 (or Terminus), and 1. Whiskers represent ±1 standard error.

Figure 1. Top: First violin part from Haydn's String Quartet in E, Op. 17/1, i, mm. 1–2. Bottom: Viewpoint representation.



	1	2	3	4	5	6	7	8	9	10	11	12
`cpitch`	64	68	71	69	66	64	68	69	71	73	64	63
`melint`	$\perp$	4	3	-2	-3	-2	4	1	2	2	-9	-1
`contour`	$\perp$	1	1	-1	-1	-1	1	1	1	1	-1	-1
`csd`	0	4	7	5	2	0	4	5	7	9	0	11

Figure 2. Top: Bar plots of the mean information content (IC) estimated for the terminal note event in the first violin (left) and cello (right) for each level of *tonic closure*. Viewpoints include `cpitch`, `melint`, `csd`, and an optimized combination called `selection`, which represents `melint` and the linked viewpoint `csd` ⊗ `cpitch`. Bottom left: Bar plot of the mean information content (IC) estimated for the terminal `vintcc` and `csdc` for each level of *tonic closure*. Bottom right: Line plot of the mean information content (IC) estimated for the combination called `composite`, which represents the dot product of $\texttt{selection}_{\mathrm{vl1}}$ and `csdc`. Whiskers represent ±1 standard error.

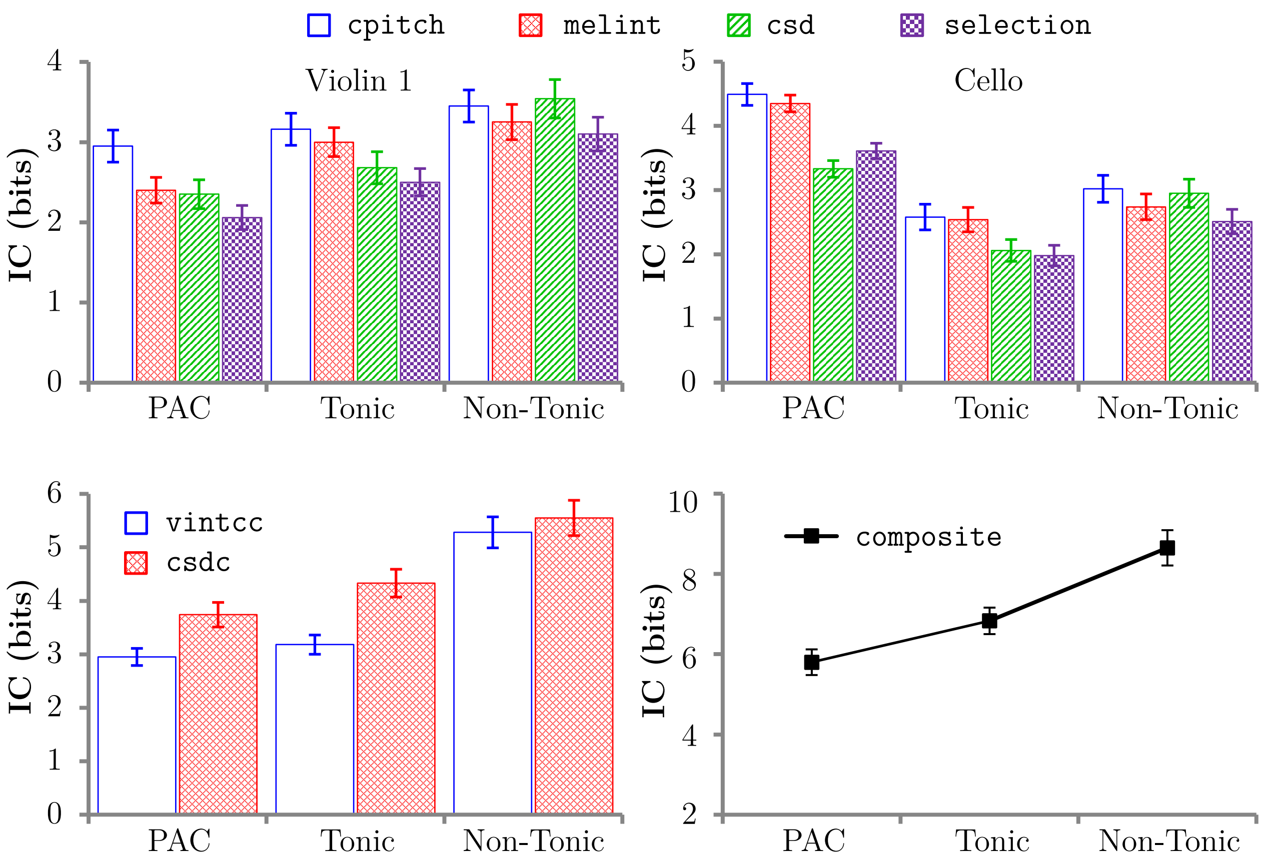


Figure 3. Top: Bar plots of the mean information content (IC) estimated for the terminal note event in the first violin (left) and cello (right) for each level of *dominant closure*. Viewpoints include `cpitch`, `melint`, `csd`, and an optimized combination called `selection`, which represents `melint` and the linked viewpoint `csd` ⊗ `cpitch`. Bottom left: Bar plot of the mean information content (IC) estimated for the terminal `vintcc` and `csdc` for each level of *dominant closure*. Bottom right: Line plot of the mean information content (IC) estimated for the combination called `composite`, which represents the dot product of `selection`$_{\text{vl1}}$ and `csdc`. Whiskers represent ±1 standard error.

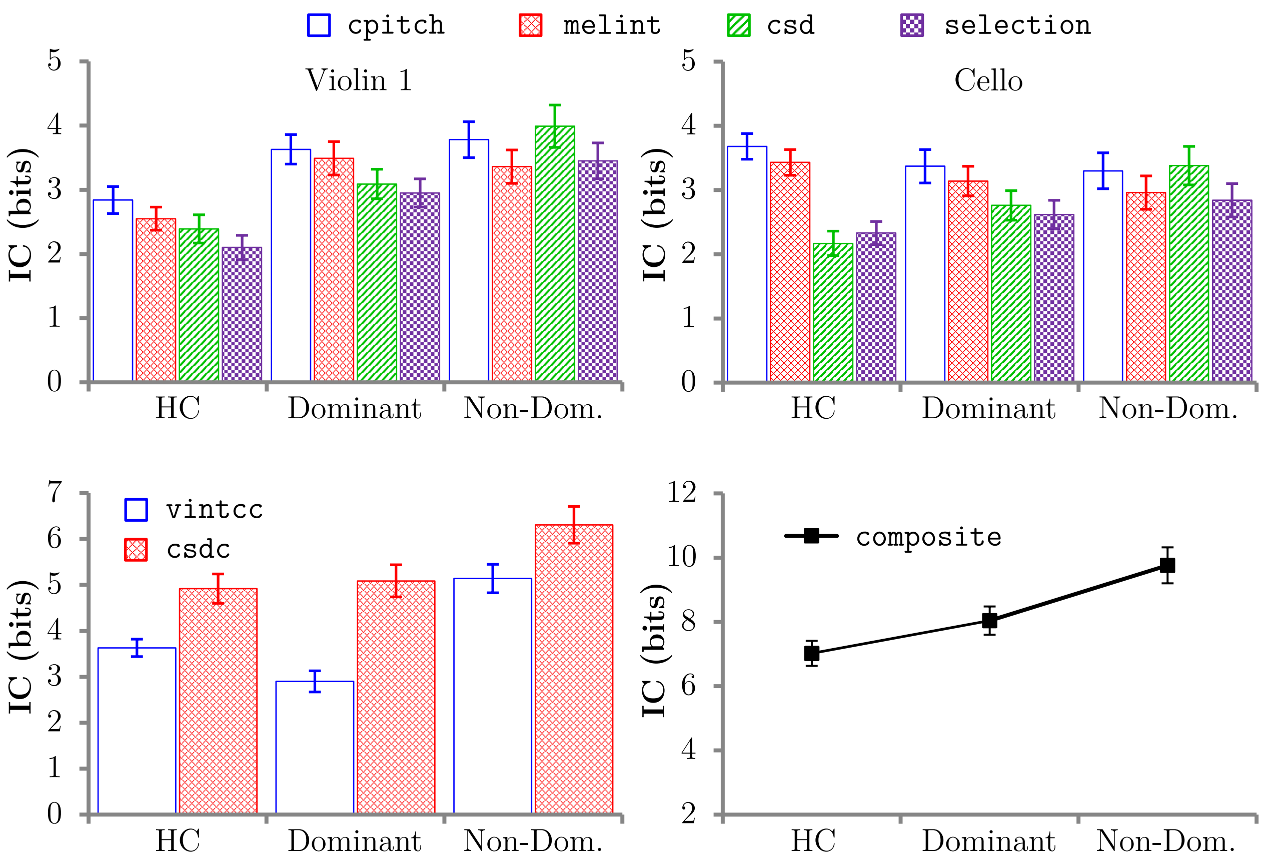


Figure 4. Line plots of the mean IC estimated for the terminal note event in the first violin (left) and cello (right) for each cadence category. Viewpoints include `cpitch`, `melint`, `csd`, and an optimized combination called `selection`, which represents `melint` and the linked viewpoint `csd` ⊗ `cpitch`. Whiskers represent ±1 standard error.

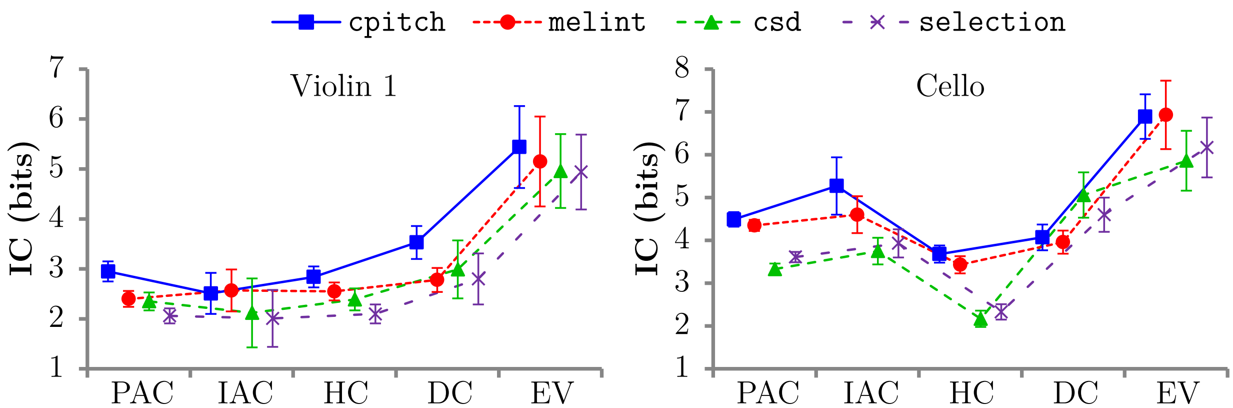


Figure 5. Left: Line plot of the mean IC estimated for the resolving chord event for each cadence category. Right: Line plot of the mean IC estimated for the combination called `composite`, which represents the product of $\texttt{selection}_{\text{vl1}}$ and `csdc`. Whiskers represent ±1 standard error.

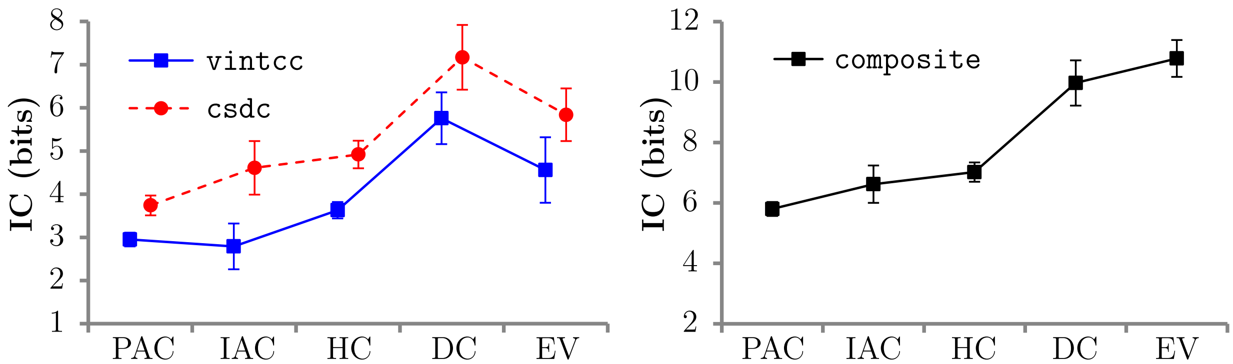


Figure 6. Time course of the mean IC estimated for the events surrounding the terminal note event in the first violin (top) and cello (bottom) for each cadence category using the viewpoint `selection`, which represents `melint` and the linked viewpoint `csd` ⊗ `cpitch`. The statistical analysis pertains to event numbers −1, 0 (or Terminus), and 1. Whiskers represent ±1 standard error.

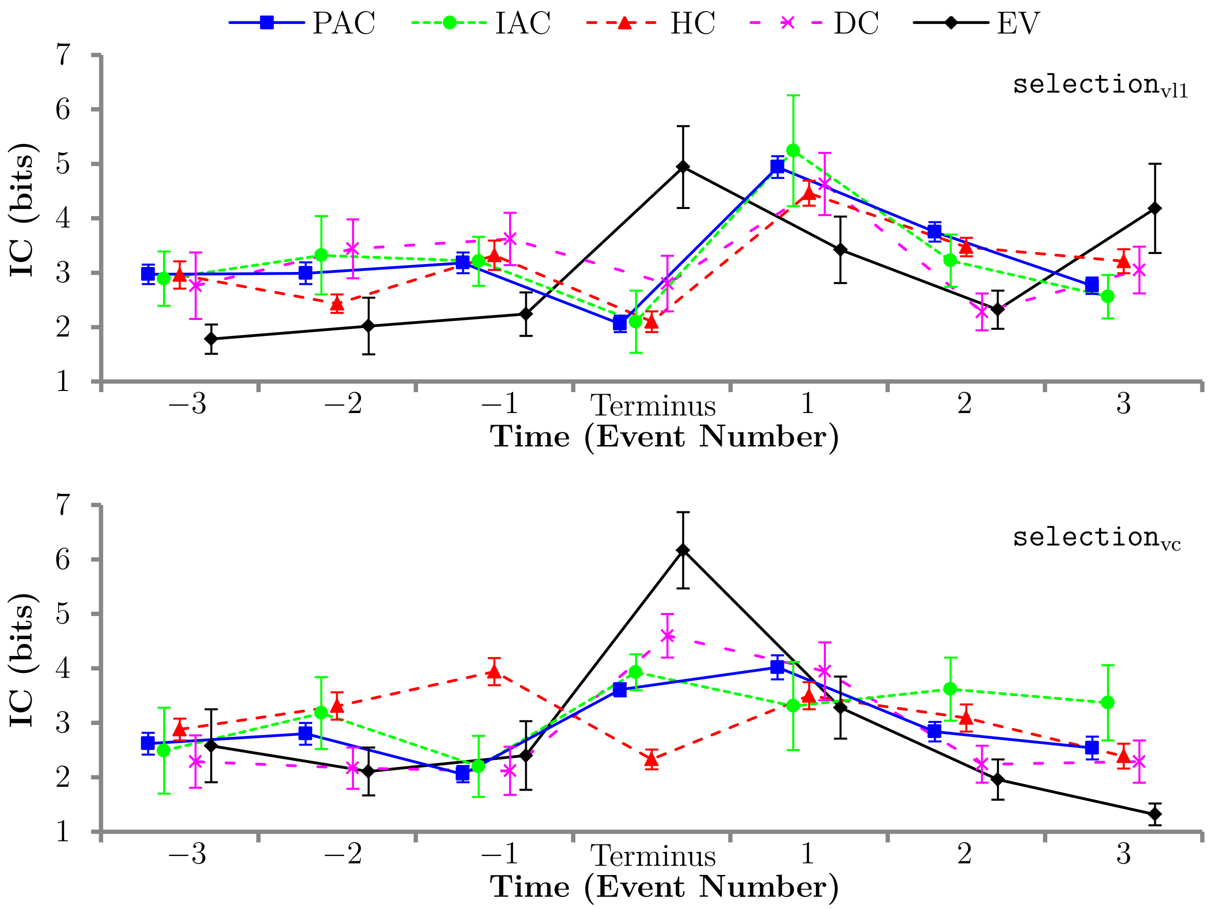


Figure 7. Time course of the mean IC estimated for the events surrounding the terminal chord event for `vintcc` (top) and `csdc` (bottom) for each cadence category. The statistical analysis pertains to event numbers −1, 0 (or Terminus), and 1. Whiskers represent ±1 standard error.