On the Relationship Between Entropy and Meaning in Music: An Exploration with Recurrent Neural Networks

Greg Cox (grcox@indiana.edu)
Department of Psychological and Brain Sciences, Indiana University
1101 E. Tenth St., Bloomington, IN 47405 USA

Abstract

Meyer (1956) postulated that meaning in music is directly related to entropy–that high entropy (uncertainty) engenders greater subjective tension, which is correlated with more meaningful musical events. Current statistical models of music are often limited to music with a single melodic line, impeding wider investigation of Meyer's hypothesis. I describe a recurrent neural network model which produces estimates of instantaneous entropy for music with multiple parts and use it to analyze a Haydn string quartet. Features found by traditional analysis to be related to tension are shown to have characteristic signatures in the model's entropy measures. Thus, an information-based approach to musical analysis can elaborate on traditional understanding of music and can shed light on the more general cognitive phenomenon of musical meaning.

Keywords: Music cognition; neural networks; information theory; entropy.

Introduction

Music is an intriguing artifact of human culture, and one of the challenges in music cognition is to explain how music is capable of having meaning for the listener. Much music carries meaning by virtue of associations to non-musical things like stories and literature (Rimsky-Korsakov's *Sheherezade*), visual imagery (Mussorgsky's *Pictures at an Exhibition*), environmental sounds (taxi horns in Gershwin's *An American in Paris*), symbols (the "cross" motif in the Fugue in C-sharp minor from Book I of J. S. Bach's *Well-Tempered Clavier*), and the meaning of text or lyrics. However, music theorists and cognitive scientists have been particularly concerned with investigating music that lacks text and that does not explicitly refer to anything non-musical.[1]

Meyer (1956) postulated that meaning in music arises from the ability of a musical event to imply or refer to other *musical* events that are expected to follow it. In a later work, he summarized his hypothesis:

> Musical meaning arises when an antecedent situation, requiring an estimate as to the probable modes of pattern continuation, produces uncertainty as to the temporal-tonal nature of the expected consequent. (Meyer, 1957, p. 416)

Within a particular style, a given musical event–e.g., a dominant chord–is expected to be followed by another musical event–e.g., a tonic chord, making for an authentic cadence. These expectations can also be violated or ambiguous–perhaps the dominant chord is followed by a submediant chord, making for a deceptive cadence. In such cases, a listener experiences tension which is manifested both in subjective reports (Krumhansl, 1996) and in physiological affective responses (Steinbeis, Koelsch, & Sloboda, 2006). Tension and its associated affective qualities–reflecting uncertainty–can thus be a signature of musical meaning.

Beyond suggesting a direct link between musical meaning and tension, Meyer's definition is readily formalized via the concept of entropy, which is a measure of both uncertainty and information content (more information is necessary to describe something that is difficult to predict). Other music theorists have made use of entropy measures in a variety of ways, including the analysis of structure in atonal music (Hiller & Fuller, 1967), stylistic variation in tonal music (Knopoff & Hutchinson, 1981), and differences between musical styles (Margulis & Beatty, 2008).

While most music theoretical studies of information in music have focused on gross properties of style or large segments of music, recently, modeling techniques from cognitive science have been brought to bear on Meyer's notion of musical meaning. Markov models and recurrent neural networks enable researchers to quantify entropy and other information measures by specifying the underlying predictive model a listener might have. Measures of information content in Markov models of music can predict structural boundaries that correspond to those assigned by human listeners to monophonic (single-part) music in the minimalist style (Potter, Wiggins, & Pearce, 2007; Abdallah & Plumbley, 2009).

However, structural boundaries are only a part of musical meaning. If meaning is related to subjective tension arising from uncertainty–i.e., entropy–it should be possible to correlate instantaneous measures of entropy (an "entropy profile") with momentary affective responses to music. For instance, an authentic cadence is a point of repose and thus should be correlated with lower entropy. A dramatic climax should be correlated with a high value of entropy (a local maximum) as it represents a large amount of tension. Human-derived entropy profiles for Bach chorale melodies (Manzara, Witten, & James, 1992) are in accord with these intuitions.

It is also likely that different dimensions of music (e.g., pitch, rhythm, harmony) contribute differently to tension and to entropy. This notion is embodied in multiple viewpoint models (Conklin & Witten, 1995), although since these models have only been applied to monophonic music, they tend to focus on pitch to the exclusion of rhythmic, harmonic, and

[1] Of course, even non-referential music is sure to remind a listener–consciously or not–of something other than the music he or she is currently hearing. However, these non-musical associations tend to vary widely between individuals and as such cannot be relied upon as a basis for musical meaning.

contrapuntal dimensions. A study of entropy as a correlate of tension should address more than just single melodic lines, since harmony and counterpoint are critical dimensions along which music can meaningfully vary.

The present study investigates the extent to which entropy can serve as a general measure of tension–and thus meaning–in music. To that end, I present a recurrent neural network as a predictive model of polyphonic (multiple-part) music and compare entropy measures derived from the model with traditional music theoretical analysis. I show that features of the traditional analysis related to subjective tension have particular signatures in the model's entropy measures, supporting the hypothesis that entropy underlies musical meaning.

A Recurrent Neural Network for Music Prediction

Recurrent neural networks (RNNs) have been fruitfully used as models of sequential prediction in many domains. In music research, they have been used to compose monophonic music both with (Mozer, 1994) and without (Todd, 1989) accompanying harmonic progressions, and to model the acquisition and perception of tonal harmony (Bharucha & Todd, 1989).

Although Markov models have seen wider–and, arguably, more productive–application in monophonic music than have RNNs, Markov models are less well suited to modeling polyphonic music. Monophonic music is easily translated into a sequence of discrete symbols drawn from a finite alphabet. It is much less clear, however, how one might translate polyphonic music into a language appropriate for a Markov model, as such music includes multiple pitch sequences updating at different rates with varying degrees of independence. For instance, to describe just the pitch transitions of a four-part piece where each part spans a diatonic octave (eight possible pitches), a naïve first-order Markov model would require a state space with $8^4 = 4096$ points and a transition matrix with $4096^2 = 16777216$ entries, and this does not even include any information about rhythm![2] Further, in any realistic training set, only a small portion of the number of possible transitions will be represented, leading to problems of over-fitting and lack of generalization (though these problems can be solved in some domains with the smoothing techniques described by Pearce & Wiggins, 2004). RNNs tend to avoid these problems, since they do not require enumerating and/or representing all state transition probabilities, but rather the weights of the network represent only those dependencies necessary to minimize prediction error. In addition, since the RNN must learn its own internal representation of the input, it will naturally converge toward representations that capture the generalities in the training set.

It should be emphasized that, as with a Markov model of music, no literal psychological reality is meant to be ascribed to the structure and training procedure of a RNN. Rather, the network should be seen only as a statistical model that mirrors the *function*–not necessarily the form–of whatever predictive model a human listener has acquired through musical experience. The resulting trained network does not–indeed, cannot–represent a listener's *entire* understanding of music in general, but is limited to representing the expectations of a listener who is familiar with a piece of music and/or its style.

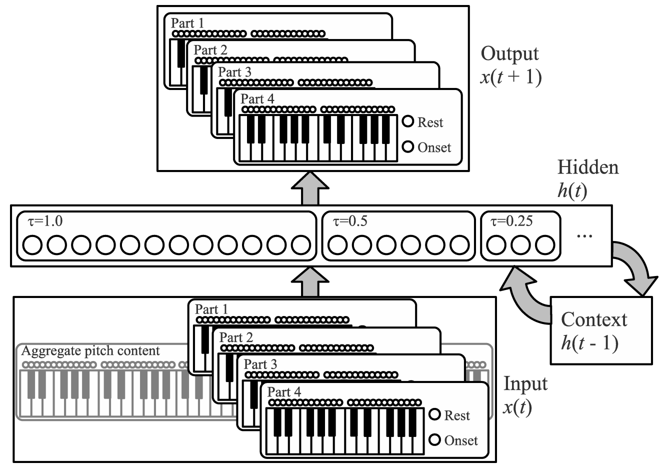

Figure 1: Schematic of the recurrent neural network used in this study, assuming 4 distinct parts. See text for details.

Architecture

The architecture of the RNN used in this study is shown in Figure 1. As in Elman (1990), the network is presented with the current state of the music, $x(t)$, and is trained–via back-propagation through time with a single time step (Rumelhart, Hinton, & Williams, 1986)–to produce the state of the music at the next time step, $x(t+1)$, at its output layer. Output layer units use a logistic activation function $f(net) = 1/(1+exp(-net))$, where *net* is the net input to the unit. Successive layers are fully interconnected.

Time Musical time is divided into discrete "time steps" of equal length, and the input, $x(t)$, describes all musical events (pitches and note onsets) occurring during that time step.

Input/Output Representation The input at time t, $x(t)$, is the concatenation of several vectors representing, for each part (i.e., distinct instrument or timbre, e.g., violin, piano, etc.) in a piece of music, the part's state along two dimensions: pitch and rhythm. The pitch dimension is represented in a localist fashion, with one unit corresponding to each absolute pitch (in twelve-tone equal temperament) that could occur in a part, including a unit representing silence or a "rest". At the input layer, a pitch unit is active (1) if it is currently sounding and zero otherwise; at the output layer, a pitch unit's activity represents the degree to which that pitch is expected at the next time step. The pitch state vector $\pi_p(t)$ for part p at time t may thus be expressed $\pi_p(t) = \langle \pi_{p0}, \pi_{p1}, \ldots, \pi_{pn}, \pi_{pREST} \rangle$ for possible pitches $0 \ldots n$ and the special *REST* "pitch". The input layer also contains a set of units representing all pitches that are sounding at the

[2] By making certain independence assumptions, it is possible to simplify a Markov model greatly, but it is not in general possible to know, *a priori*, what those assumptions should be.

current time across all parts, to allow for generalization of pitch content between parts. However, the network is only trained to predict the pitches of each part individually, not this aggregate pitch content.

An additional unit for each part represents its state along the rhythmic dimension: this unit is active (1) when the current time step contains a note onset within that part, and is otherwise inactive (when the part is silent or sustaining a previous pitch). At the output layer, this unit can be interpreted as the probability $\rho_p(t+1)$ that part p will contain a note onset at at time $t+1$. Note that the assumption of independence between rhythm and pitch in the input/output representation permits the analysis of each component separately. However, independence of representation does not guarantee probabilistic independence, as both pitch and rhythm units are treated equally in the network's internal representation in the hidden layer.

Hidden Layer Hidden unit activations are a function of the current input, the hidden layer at the previous time step (also called the "context" layer), and each unit's own prior activation. The activation of hidden unit h_i at time t is

$$h_i(t) = \tau_i f\left(\Sigma_j w_{ij} x_j(t) + \Sigma_k w_{ik} h_k(t-1)\right) + (1-\tau_i) h_i(t-1),$$

where $f(\cdot)$ is the logistic activation function described above, w_{ij} is the weight from input unit j to hidden unit i and w_{ik} is the weight from context unit k to hidden unit i. The different time constants τ_i cause the hidden units to change at varying rates over time, permitting the representation of multiple time scales at the hidden layer (Mozer, 1992).

For simplicity, I assume that the number of hidden units with time constant τ is $N_\tau = \lfloor \tau N_1 \rfloor$ where N_1 is the number of units with $\tau = 1$ and $\lfloor \cdot \rfloor$ is the floor function, ensuring that there will be only a finite number of hidden units and time scales represented. In the simulations reported here, each τ is the reciprocal of either a power of 2 or a power of 3, i.e., $\tau = 2^{-\gamma}$ or $\tau = 3^{-\gamma}$ for $\gamma = 0, 1, 2, \ldots$. The choice of time constant scales based on 2 and 3 derives from the predominant metrical subdivisions (duple and triple meters) in Western music, which is the domain of the current study. Thus, the hidden layer best represents information at time scales that are likely to be most salient.

Measures of Entropy

Although there are many ways to measure entropy within the current modeling architecture of the RNN, I will focus on four simple measures, three of which are used in the subsequent musical analyses. For any part p, the pattern of activity over its pitch units (including the "rest" pitch) at the output layer, π_p, can be normalized to sum to one, such that it can be considered a probability distribution over pitches. Then, the entropy with regard to pitch in part p at time t is $H_p^{pitch}(t) = -\sum_{i=0}^{n,REST}\left[\pi_{pi} log_{(n+1)}(\pi_{pi})\right]$, where the base of the logarithm normalizes the entropy to the range from zero to one. Similarly, the entropy with regard to rhythm in part p at time t is $H_p^{rhythm}(t) = -\rho_p(t) log_2(\rho_p(t)) - (1-\rho_p(t)) log_2(1-\rho_p(t))$.

To measure entropy over the entire ensemble rather than within each part, an aggregate pitch probability vector $\pi^* = \langle \pi_0^*, \pi_1^*, \ldots, \pi_n^* \rangle$ is created, where $\pi_i^* = C \sum_{p=0}^{P} \pi_{pi}$, i.e., the sum of the probability assigned to pitch i by each of the P parts, normalized (by constant C) to sum to one. The entropy of π^* can then be computed. The rhythmic entropy of the ensemble is computed over the joint distribution of the onset probabilities of each part. Pitch entropy represents uncertainty about *what* pitches will occur, while rhythmic entropy represents uncertainty about *when* those pitches will occur.

Long-Term and Short-Term Models

As in work with multiple viewpoint models of music (Conklin & Witten, 1995), for each piece of music to be analyzed, two of the above-described networks are trained. The first network is trained on a representative sample of a particular style of music and is meant to represent more global stylistic characteristics acquired by the listener over a longer time span, hence it is called the long-term model (LTM). A second network is trained on just a single piece of that style and is meant to represent knowledge of that piece in particular acquired over less time, hence it is called the short-term model (STM). This distinction is akin to that between "schematic" (LTM) and "veridical" (STM) knowledge made in Justus and Bharucha (2001). Both models produce patterns of activation over output units representing the expected pitch and rhythmic state of each part. These patterns can be combined to form an aggregate prediction from both the STM and LTM models[3]. Following Pearce, Conklin, and Wiggins (2005), this combination is a weighted geometric mean of the output activations for each dimension of each part of each model, where the weight is inversely proportional to the entropy of the activity over the relevant dimension of each part. For example, the aggregate activation of pitch π_i in part p (aggregate rhythm activation is analogous) would be

$$\bar{\pi}_{pi} = \left[\left(\pi_{pi}^{STM}\right)^{\frac{1}{H_{STM}^{pitch}}} \left(\pi_{pi}^{LTM}\right)^{\frac{1}{H_{LTM}^{pitch}}} \right]^{\frac{1}{\frac{1}{H_{STM}^{pitch}} + \frac{1}{H_{LTM}^{pitch}}}} .$$

The effect of combining the STM and LTM in this way is to emphasize "points of agreement" between them. For example, if they both strongly predict a particular pitch, the aggregate activity ascribed to that pitch will be very high. If one model is ambivalent (high entropy) while the other is certain (low entropy) of a particular pitch, the aggregate activity will accrue to the pitch of which one model is certain. If both models are certain but disagree, activity will be diffused over all possible pitches, leading to high entropy of the aggregate STM-LTM prediction.

[3] Justus and Bharucha (2001) found that schematic (LTM) and veridical (STM) knowledge made independent contributions to musical expectations; their results are consistent with a weighted geometric mean of those two sources of information.

Applying the Network: Haydn's String Quartet Op. 20, No. 3, First Movement

Because Markov models are already well-suited to modeling monophonic music and RNNs have already been shown to deal well with monophonic melodies, even those with accompanying harmonic progressions, I wanted to explore polyphonic music that did not have a simple "melody plus chords" texture–in other words, music that has been difficult to model with previous approaches. There is also an inherent difficulty in correlating entropy with tension, since tension in a listener is not directly observable. As such, I will consider tension as it is normatively described by traditional music theoretical analysis. The analytical procedure described below has been replicated with a variety of corpora, including Bach chorales, Chopin piano preludes, and Schönberg's *Pierrot Lunaire*, with similar results regarding the relationship between entropy and traditional accounts of tension. To show how an analysis of entropy relates to traditional approaches, I report here the results of a single analysis in detail.

The Op. 20 string quartets of Joseph Haydn share many stylistic characteristics–for example the use of "sonata form", a typical classical dramatic form, in the first movement of each quartet. Yet despite the regularities among the quartets and between their first movements in particular, they contain many deviations from standard practice. Both global regularities and local idiosyncrasies contribute to the dramatic content of these pieces and make them prime targets for analysis.

The third quartet, in G minor, is particularly dramatic, containing prolonged periods of tension, metrical ambiguity, and various surprising moments. I used the above-described RNN model to calculate measures of entropy for each time step in the first movement of this quartet. I then compared these measures to features derived from a music theoretical analysis of the piece in terms of its formal and dramatic structure.

Training

All pieces on which the RNN were trained were encoded as MIDI files, with each instrument (two violins, viola, and cello) assigned to a different part and thus separately represented in the RNN's input and output layers. In total, 247 units were used to represent the input (pitch and rhythm units for all four parts separately, as well as the aggregate pitch content) and 177 units were used in the hidden layer. The back-propagation learning rate parameter was set at 0.0625 and time steps were set at sixteenth-note duration.

The LTM was trained on the first movements of all six quartets in Op. 20 (19006 total time steps). All pieces of the training set were transposed to either C-major or C-minor as appropriate to eliminate effects of absolute pitch (since the model uses a localist pitch representation). The LTM was trained in cycles, during each of which it was trained on all six training pieces in random order. Training continued until mean accuracy–defined as the mean probability assigned to each time step in the training pieces–did not change by more than 0.0001 for 10 consecutive cycles. In all, the LTM was trained for 2000 cycles and achieved a final accuracy over the entire training set of 0.276 (range: 0.179 to 0.383).

The STM was trained on only the first movement of Op. 20, No. 3 (4332 total time steps). Using the same stopping criterion, the STM was presented with this movement 2500 times and achieved a final accuracy of 0.751. The combined LTM and STM models, which produced the output analyzed below, achieved an accuracy of 0.456 on the movement.

Simulations were also conducted which varies the number of hidden units, learning rate, and size of the LTM training corpus (for example, by including a wider selection of Haydn string quartet movements from Op. 17). The only major effect of these variations was that accuracy was improved with additional hidden units, but the form of the entropy profiles remained the same; specifically, major points of inflection were all at the same place and in the same direction.

Analysis of Entropy Profiles

The pitch and rhythmic entropy profiles derived from the combined STM and LTM are shown in Figure 2. Only the first repeat of the exposition (the first section of a sonata form piece; mm. 1-94) is shown, as this will be the focus of the subsequent analysis. Lacking a principled method of integrating pitch and rhythmic entropy, they are here considered separately, although both are assumed to contribute to a listener's subjective sense of tension. To enable the analysis of trends in the entropy measures, they were smoothed by convolving the raw entropy measures with an exponentially-decaying impulse response filter with weights $\psi(t) = e^{-\lambda t}$, where decay constant $\lambda = \frac{1}{32}$ corresponds to a mean lifetime of four measures (32 time-steps). Thus, the values shown in Figure 2 represent a "memory" of the instantaneous entropy that emphasizes the last four measures. The following analysis owes much to the work of Drabkin (1999), particularly pp. 105-111. Additional analytical material may be found in Grave and Grave (2006), especially pp. 190-192.

The only perfect authentic cadence in the exposition occurs at the end of the first phrase in m. 7, where there is a clear local minimum in pitch entropy as well as a low plateau in rhythmic entropy[4]. Mm. 8-26 effect a modulation from the home key of G minor to its relative major, B♭, all the while increasing the tension for a strong resolution to a B♭ harmony. This increase in tension is mirrored by increasing pitch and rhythmic entropy, where pitch entropy reaches a local maximum on the second beat of m. 24 with the introduction of a novel unison figure that prolongs the tension until the B♭ resolution in m. 27.

The second theme group (mm. 27-40) maintains a consistent pitch entropy while rhythmic entropy builds until the cello's eighth-note pulse disappears in m. 34, leaving just a high violin melody with the other instruments holding chords in long rhythmic values. The decrease in rhythmic entropy is

[4] In simulations with Bach chorales (not reported here), resolutions of authentic cadences also correspond to local minima in entropy measures while deceptive cadences produce no change or an increase in entropy.

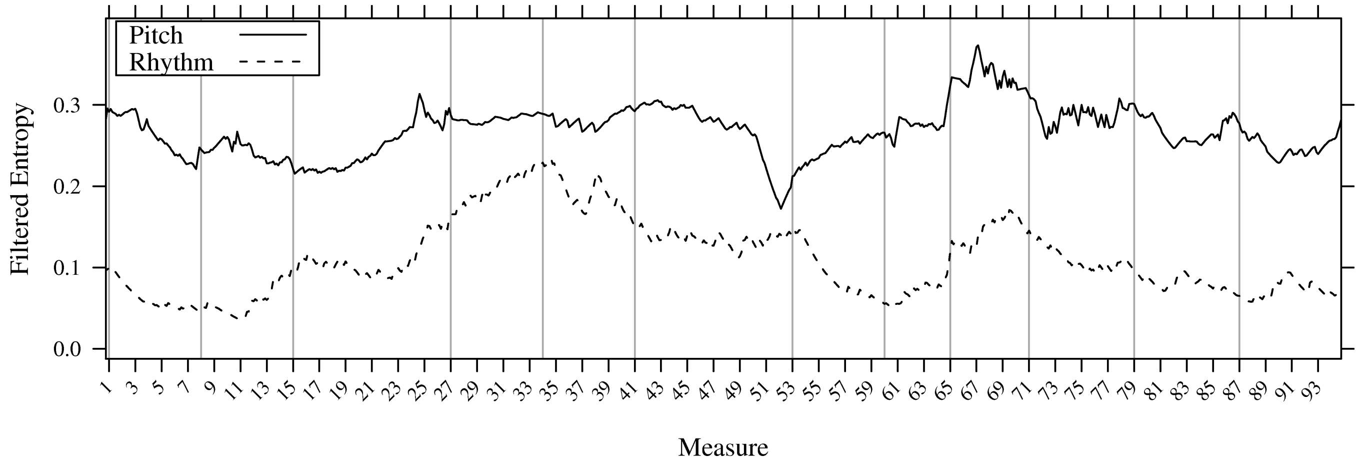

Figure 2: Ensemble pitch and rhythm entropy profiles for the exposition (mm. 1-94) of Haydn Op. 20, No. 3, first movement.

interrupted in mm. 37-38, when pitch changes become staggered between the instruments. Mm. 41-52 are more transitory and fragmented, with a large drop in pitch entropy during the violin solo in mm. 50-51 (greater certainty arising from predicting fewer separate parts). The long, regular rhythmic durations of mm. 53-59 continue the drop in rhythmic entropy while the chromatic harmonies increase pitch entropy until a break is reached at a deceptive cadence in m. 60.

This is followed by an F-major statement in mm. 61-64, then in mm. 65-66 by an "utter non sequitur–a fortissimo fanfare, poised on a first-inversion B♭ triad, with no compelling relationship to the immediately preceding or following material" (Grave & Grave, 2006, p. 190). This surprising event is naturally accompanied by a spike in both rhythmic and pitch entropy. Contrary to what might be implied by the B♭ fanfare–strong thematic material emphasizing the new key of B♭–we are instead treated in mm. 67-70 to the opposite: a softer, tonally ambiguous reprise of mm. 61-64. Mm. 67-70 are at a softer dynamic, played by solo violin instead of the entire quartet, and in a more restricted and chromatic melodic range. This unexpected consequent is assigned the highest pitch entropy in the entire exposition.

Rhythmic entropy continues to build until a resting point is reached at m. 70 on an unclear tonality. The succeeding violin solo and its accompaniment in mm. 71-77 is metrically ambiguous, implying a triple meter when in fact the duple meter still prevails. In this instance, the gradually diminishing ensemble rhythmic entropy is not in accord with this ambiguity, which should result in a higher rhythmic entropy for this passage. However, the rhythmic entropy of the individual parts, shown averaged in Figure 3, does show the expected staggered increase from mm. 71-77.

The remainder of the exposition is on more solid tonal and metrical footing. Of particular interest is the jump in pitch entropy in mm. 85-86, corresponding to another instance of the unison figure from mm. 24-25 and serving the same purpose–to prolong tension before before reaching a harmonic resolution–and producing the same effect on the entropy profile–an increase in pitch entropy whilst rhythmic

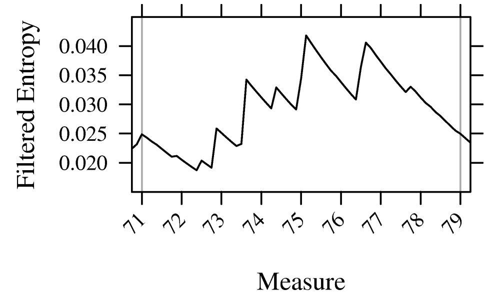

Figure 3: Rhythmic entropy averaged between parts for mm. 71-79 of Haydn Op. 20, No. 3, first movement.

entropy is unaffected. Although rhythmic entropy reaches a local maximum at m. 90 and begins to fall after a constant eighth-note pulse is established in the violins, pitch entropy increases toward the end of the exposition, reflecting the fact that the end of the exposition can be followed by either a repeat of the exposition (return to m. 1) or the start of the next section. In both cases, the rhythmic surface is the same, but the pitches are different and are assigned to different instruments, thus it is logical that there would be more uncertainty about pitch than rhythm at the end of the exposition.

Discussion

Analyses like the one presented above show that entropy derived from a predictive model of music can correspond to dramatically important features of music. Specifically, the entropy measures employed are sensitive to the calming effect of cadences (m. 7), the build-up of tension prior to resolutions (mm. 8-26), differential effects of textural change (mm. 27-60), and the shocking effects of interruptions (mm. 24-25, 85-86) and their consequents (mm. 65-70). Because a listener's subjective sense of tension is also affected by these features, this suggests a relationship between entropy and tension–and thus, perhaps, to musical meaning.

It is, perhaps, remarkable that such a relationship may be found, given the limitations of the current model. The model includes no information about dynamics, timbre, and expressive timing. A more realistic pitch representation, while increasing the model's complexity, might also improve its performance (Mozer, 1994). Further, the use of a RNN at all imposes severe limitations on the approach outlined in this paper. While RNNs enable the analysis of music that is not amenable to other modeling techniques, they are slow to train, limited in the size of the corpus on which they can be trained, and, in the form presented here, cannot generalize to other ensemble types. The application of computational cognitive models to music is still in its infancy, and future research is sure to improve upon the techniques explored thus far. Future work must also compare model-derived entropy measures with human tension judgments (as in Krumhansl, 1996). This will elaborate on the relationship between entropy and tension, including the contributions of different sources of uncertainty (e.g., pitch and rhythm) to overall tension.

Even given the limited state of our current knowledge, it is possible to show that meaningful musical features correlate with features of musical entropy, given an appropriate predictive model. If human listeners have a similar predictive model "in mind"–consciously or not–as they listen to music, this provides great insight into the nature of music cognition and creation. The reasons why certain patterns recur within a style and that listeners have consistent responses to those patterns and violations thereof are not arbitrary–they can be understood in terms of prediction and uncertainty. With the advent of formal cognitive models, we can leverage this principle to better understand music that resists conventional analysis, for example, styles with few examples (e.g., the oeuvre of many idiosyncratic modern composers) or for which there is insufficient access to primary sources (e.g., historical and ethnomusicological studies). While more sophisticated methods will allow us to better elucidate the nature of entropy in music, it is clear that Meyer's (1956) thesis is still a viable approach to understanding the nature of meaning in music.

Acknowledgments

The author wishes to thank Peter Fontana, James Fry, Dora Hanninen, Michael Jones, and Richard Shiffrin for their guidance, suggestions, and support, as well as several anonymous reviewers.

References

Abdallah, S., & Plumbley, M. (2009). Information dynamics: Patterns of expectation and surprise in the perception of music. *Connection Science*, *21*(2), 89–117.

Bharucha, J. J., & Todd, P. M. (1989). Modeling the perception of tonal structure with neural nets. *Computer Music Journal*, *13*(4), 44–53.

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

Drabkin, W. (1999). *A reader's guide to Haydn's early string quartets*. Westport, CT: Greenwood Press.

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

Grave, F., & Grave, M. (2006). *The string quartets of Joseph Haydn*. New York: Oxford University Press.

Hiller, L., & Fuller, R. (1967). Structure and information in Webern's Symphonie, Op. 21. *Journal of Music Theory*, *11*(1), 60–115.

Justus, T. C., & Bharucha, J. J. (2001). Modularity in musical processing: The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 1000–1011.

Knopoff, L., & Hutchinson, W. (1981). Information theory for musical continua. *Journal of Music Theory*, *25*(1), 17–44.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's piano sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, *13*(3), 401–432.

Manzara, L. C., Witten, I. H., & James, M. (1992). On the entropy of music: An experiment with Bach chorale melodies. *Leonardo Music Journal*, *2*(1), 81–88.

Margulis, E. H., & Beatty, A. P. (2008). Musical style, psychoaesthetics, and prospects for entropy as an analytic tool. *Computer Music Journal*, *32*(4), 64–78.

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

Meyer, L. B. (1957). Meaning in music and information theory. *The Journal of Aesthetics and Art Criticism*, *15*(4), 412–424.

Mozer, M. C. (1992). Induction of multiscale temporal structure. In J. E. Moody, S. J. Hanson, & R. P. Lippmann (Eds.), *Advances in neural information processing systems IV* (pp. 275–282). San Mateo, CA: Morgan Kaufmann.

Mozer, M. C. (1994). Neural network music composition by prediction: Exploring the benefits of psychoacoustic constraints and multi-scale processing. *Connection Science*, *6*, 247–280.

Pearce, M. T., Conklin, D., & Wiggins, G. A. (2005). Methods for combining statistical models of music. In *Computer music modeling and retrieval.* Berlin / Heidelberg: Springer.

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

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), 295–322.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & PDP Research Group (Eds.), *PDP.* Cambridge, MA: The MIT Press.

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380–1393.

Todd, P. M. (1989). A connectionist approach to algorithmic composition. *Computer Music Journal*, *13*(4), 27–43.