Music, Expectation, and Information Theory

D. Gareth Loy

> *Music conveys general forms of feelings, related to specific ones as algebraic expressions are related to arithmetic [expressions].*
> Susanne Langer [5]

Abstract What makes a musical work successful? In Darwinian terms, music is successful if listeners attend to it, repeatedly, for then it can live on. However, attention is fleeting: successful music holds listeners' interest by manipulating their expectations using deception and confirmation. The ratio of the rate at which listeners follow music to the rate at which music unfolds is a predictor for musical success. This paper informally presents a theory of musical interest, based on some ideas from music theory, cognitive psychology, and information theory.

1   When is Music Successful?

Music is successful in Darwinian terms if we are repeatedly willing to hear it. Successful music is so because it cultivates and sustains listeners' interest. In no other way can music live on. Music's most powerful attractant is our *curiosity*. By exploiting it, successful music lives to be heard another day.

We are curious when we want to learn about something. When the discovery process is going well, the learner is *engaged*. This is more than simple attraction. If the discovery process continues commensurate with the rate of new information received, we can sustain our curiosity if we wish. The same is true of music: we can, if we wish, remain engaged if the rate at which we follow the music is commensurate with the rate at which it unfolds. However, if the rates are unmatched, and our minds outrace or fall behind the music, we lose interest.

D.G. Loy (✉)
Gareth Inc., POB 151185, San Rafael, CA 94915, USA
e-mail: dgl@garethinc.com

If our minds race ahead—figuring out where the music is going faster than it gets there—we risk boredom, and loose interest. After all, it is redundant, and a waste of time, for us to know what will happen *and* to bother hearing it out. On the other hand, if we fall behind because the music outstrips our ability to keep up, we grow frustrated and lose interest.

Interest in music is closely tied to the rate at which we can make sense of what we hear. To follow music means to be able to orient oneself, to understand what has been heard, and to have a prediction, or an expectation, of where it is going. If our understanding increases commensurate with the rate of musical information then we believe ourselves to be in possession of enough knowledge to remain current as it unfolds, and to have some confidence that we can anticipate forthcoming musical events. Some degree of such confidence is required for interest to persist. But it is the *vulnerability* of this confidence that successful music exploits.

Even for very simple music, we form and evaluate large numbers of mostly unconscious predictions as we listen to music. The key to engaging listeners is to satisfy some expectations while frustrating others as the music unfolds. This is the art of *entertainment*.

Example of musical expectation Figure 1 shows an elementary motive of four notes sequenced up repeatedly by diatonic steps.

Suppose you were hearing it played for the first time. By the end of measure 2, you'd probably have heard the repeated motive. You might think, "I bet the music is sequencing a four note motive up diatonically." If, as in the third measure, the music meets your expectation, your prediction is confirmed [4]. You feel a fleeting sense of satisfaction... *and, curiously, the music starts to lose your interest* because no sooner is the pattern you've predicted realized than it ceases to be interesting: because there is little to no new information to digest, it's a waste of time to continue listening.

If the musical pattern continues unvarying into the fourth measure as shown, a new sensation, *boredom* may arise. Interest is allergic to deadeningly predictable patterns. Music dies when listeners don't care to hear it. But suppose instead the music veers off as shown in Fig. 2.

Here, after exactly 2.5 repetitions of the four-note motive, the music switches from horizontal to vertical motion—from melodic sequencing to a dominant-tonic (V–I)

Fig. 1 Elementary sequenced motive

Fig. 2 Elementary sequenced motive with cadence

cadence. The listener, having already heard two repetitions of the motive, expects the pattern to continue and is surprised by its interruption. Surprise is invoked by the introduction of an asymmetry that violates the listener's expectations, and the surprise serves to fetch the listener's interest back, thereby *entertaining* the listener.

Music requires a degree of structural ambiguity to gain and maintain interest. The structure of successful music must continually mutate to sustain listeners' engagement, i.e., to entertain listeners. Violating motivic regularity is but one way to accomplish this.

How is it that we were able to predict the evolution of the musical motive in Fig. 2 even before we'd heard it all the way through? This suggests we carry models—schema—of what we expect, which we apply to fathom novel circumstances. Schemas describe patterns of thought that organize and categorize our experiences, and express the relationships among them.

Aristoxenus said,

> Musical cognition implies the simultaneous recognition of a permanent and a changeable element... for the apprehension of music depends upon those two faculties, sense perception and memory; for we must perceive the sound that is present, and remember that which is past. In no other way can we follow the phenomenon of music. – Aristoxenus [1]

How indeed can we follow music unless we can compare the sound that arises to what we expected to hear? Leonard Meyer said,

> Emotion or affect is aroused when a tendency to respond is arrested or inhibited... What a musical stimulus or a series of stimuli indicates... [is] not extramusical concepts and objects but other musical events which are about to happen... Embodied musical meaning is, in short, a product of expectation. – Leonard Meyer [6]

Representational momentum When comparing what we hear in the present to our expectations from the past, we experience varying degrees of confirmation and surprise, much as, when following a ball in flight, we may experience confirmation if it hits its mark, and surprise if it is suddenly deflected. Freyd and Finke discovered that,

> Under appropriate conditions an observer's memory for the final position of an abruptly halted object is distorted in the direction of the represented motion, much as a physical object continues along its path of motion because of inertia [2].

The authors termed this phenomenon *representational momentum* [3].

We can adapt the concept for musical purposes by reference to Fig. 2, where the repetitive motivic sequence sets up representational momentum in the listener's mind in the form of a belief that the pattern will continue. The surprise elicited when the cadence breaks the pattern is analogous to the surprise that would be elicited by the "abruptly halted object" referenced by Freyd and Finke. Surprise demonstrates the presence of the representational momentum in the listener's mind, for there would be no surprise were there no expectation that the phenomenon—either the ball flying through the air, or the melody sequencing—would continue.
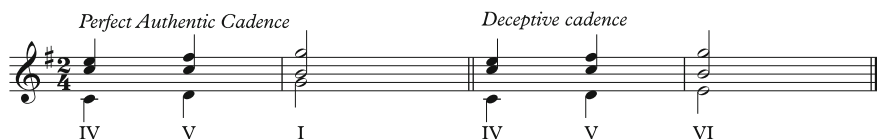
Fig. 3 Perfect authentic cadence; deceptive cadence

Representational momentum and the deceptive cadence The canonical finish to a musical phrase, the perfect authentic cadence, shown in the first two measures of Fig. 3, outlines the chordal sequence from the subdominant chord (IV), to the dominant (V), finally resolving to the tonic (I) chord. If completed, the listener expects a full stop to the musical phrase in progress. The music may go on, but one musical idea has stopped and another has begun.

The *deceptive cadence* (Fig. 3) subverts the listener's expectation of phrase completion. It begins like the perfect authentic cadence, but at the last chord, it "resolves" to the VI chord instead of the I. The triad on VI shares two of its three degrees with the tonic I triad, so the VI triad mimics the tonic enough so that the ear is not completely derailed by its substitution for the tonic. However, it is not the tonic, and until that moment, the listener expected resolution to the I chord, and is surprised when the VI is substituted, reengaging the listener's interest.

The deceptive cadence is the musical equivalent of "bait-and-switch", whereby what we are expecting is not what we get. Imagine you are a hunter in the woods and are about to bag a fat squirrel for dinner, but it slips away. This is the effect of the deceptive cadence on the ear. The listener is now more "hungry" for the proper cadence; the composer can now build up to a more charged climax.

In order to eat, the hunter must continue hunting after missing the squirrel; just so, after a deceptive cadence, the listener must continue to seek resolution. Composers use this to extend the duration of a musical phrase. Figure 4 shows a deceptive cadence and its continuation in the opening of the second movement of Mozart's *Piano Sonata in C*.

To the listener, the *meaning* of the deceptive cadence (using Meyer's definition) is that there is more to the current phrase that is still to come.
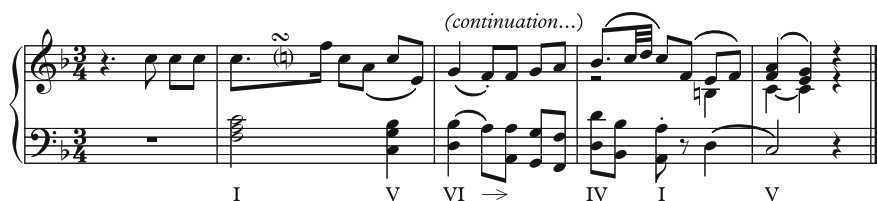

Fig. 4 Mozart *Piano Sonata in C*, K. 330, opening of 2nd movement

2   Information Theory

In 1928, Harry Nyquist proposed that a signal must be sampled at twice its highest frequency so as to have enough information to completely reconstruct the original signal from its sampled representation [7]. Therefore, a signalling system with bandwidth B has a maximum data rate $2B$. A transmission system having K distinct amplitude levels represented with binary encoded values has a maximum data rate D of:

$$D = 2B \log_2 K. \tag{1}$$

Shannon and Weaver [9] extended Nyquist to account for noise:

$$C = B \log_2 (1 + S/N) \tag{2}$$

where: C = channel capacity (bits/second), B = hardware bandwidth, S = average signal power, N = average noise power, S/N is signal-to-noise ratio.

The channel capacity C required to send a signal depends upon its degree of regularity. If a signal is highly ordered or predictable, it has a high degree of *redundancy*, and a summary of the redundant components of the signal can be transmitted instead of the entire signal, requiring less channel capacity C. If a signal is highly unordered or unpredictable, it has a high degree of *entropy*. The higher the degree of entropy, the fewer of its components are redundant. Components that cannot be summarized must all be transmitted, requiring more channel capacity C.

Information theory borrowed the term entropy from chemistry, where entropy is the thermodynamic probability of a molecular system, that is, a measure of the ways in which the *energy* of a molecular system is distributed among the possible motions of its particles. In information theory, entropy is a measure of the ways in which the *information* of a signaling system is distributed among its possible communications [8].

Surprisal is a measure of the uncertainty in a communication. Surprisal is analogous to the experience of "surprise", and it relates to the probability of an expected outcome.

Probability ranges over the unsigned unit interval (0.0–1.0) where, for probability $p = 1.0$ corresponds to absolute certainty, $p = 0.0$ corresponds to absolute uncertainty. Classically, probability values are defined for all time—they do not change.

Surprisal is inversely related to probability. In the limit as the probability of an event goes from 1.0 to 0.0, surprisal goes from zero to infinity. That is, for surprisal s and probability $p = 1 \rightarrow s = 0$, $p = 0 \rightarrow s = \infty$.

If an event will occur no matter what ($p = 1$), then there is no surprisal. For example, a coin toss will be either heads or tails—no surprise there. On the other hand, if there is a vanishingly small probability that an event will occur, then the

surprisal goes to infinity. For example, suppose you win the lottery—your surprise knows no bounds! Therefore, $p = \frac{1}{2^s}$. Solving for s:

$$s = \log_2 \frac{1}{p} = -\log_2 p = -\frac{\ln p}{\ln 2}. \tag{3}$$

Surprisal is the inverse log probability of a token appearing in a message. Surprisal s relates to the bandwidth required to communicate a particular message that has probability p.

Frequency and surprisal The frequency of probable events has an amplifying effect on expectation. Suppose you randomly find a dollar on the sidewalk one day: you are surprised. But if you randomly find a dollar on the sidewalk several days in a week, you are astonished! In information theory, frequency is how often a token appears in a message.

If there are N tokens in message X and the ith token occurs N_i times, then $\frac{N_i}{N}$ is its frequency.

Average surprisal The average surprisal of a message is the normalized sum of the expectancy of its tokens. In music, the surprisal of a melody is the normalized sum of the expectancy of its notes.

For example, let all the keys on a piano be independently played. Let each piano key be $k_i, i = 1, 2, 3, \ldots, M$, where M is the number of keys. If N notes can arise in a melody X, then its average surprisal H is:

$$H(X) = \frac{1}{N} \sum_{i=1}^{M} \frac{N_i}{N} s_i \tag{4}$$

We normalize the sum by the number of tokens in the message to facilitate comparing surprisal across messages of varying length.

Examples of surprisal Let us take the hypothesis that the keys near middle-C are most frequently played on the 88-key piano keyboard. The normal (Gaussian) probability distribution function with mean $\mu = 44$ (corresponding to the center key of the keyboard, which has the pitch E4, that is, the pitch E above middle-C) and standard deviation $\sigma = 1$ is shown in Fig. 5. The corresponding normalized average surprisal is shown in Fig. 6.

If the hypothesis is correct, then we should expect to hear the keys near the center of the keyboard played most frequently on the piano, and if our expectation is violated, we are surprised. Thus, by Eq. 4 we would be surprised by a melody played entirely by high and low keys, and little surprised by a melody played near the center of the keyboard.

Taking the average surprisal function shown in Fig. 6, we can calculate the surprisal of various melodies played on the piano, as follows. The melody of Antonio Carlos Jobim's *One Note Samba* is sung on a single note. ("Eis aqui este sambinha, feito numa nota só ...") If the melody is played on E4, then the average surprisal

Fig. 5 Probability density function
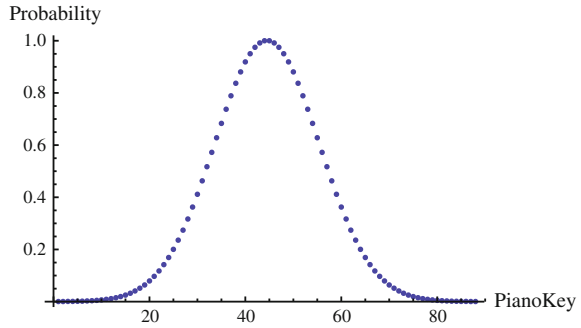
Fig. 6 Corresponding surprisal
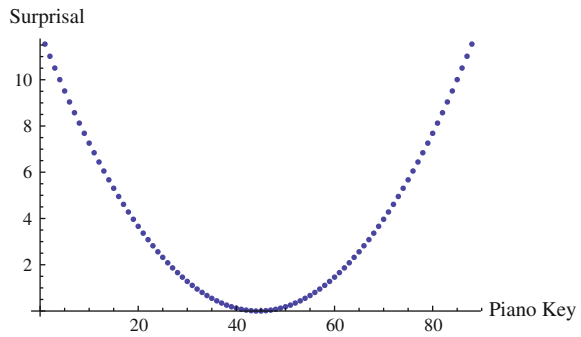
of the first 32 notes of this melody is 0. The average surprisal of a chromatic scale played in the middle of the piano keyboard would be very low, on the order of 0.006. The average surprisal of a chromatic scale far from the center of the keyboard would be higher, on the order of 0.8. The average surprisal of a random 12-note melody would be about 0.33.

Clearly, the meaningfulness of surprisal depends on the validity of the hypothesis. The relevance of information theory to music is its formalization of expectation and surprisal; but its ultimate usefulness to music theory depends upon the development of a corpus of theories that correctly capture the actual experience of listeners. It is not clear that this is possible to do in absolute terms. Given the evident variety of music around the world and through time, one assumes that the relevant musical schema depend upon a highly contextual field of cultural antecedents that are difficult to elicit, let alone classify.

Uncertainty As the total number of events in a message N increases to infinity, the event frequency $\frac{N_i}{N}$ tends to its static probability p_i. By combining Eq. 4 with the definition for surprisal s_i (Eq. 3) and substituting p_i for N_i/N, we have:

$$H(X) = -K \sum_{I=1}^{M} p_i \log_2 p_i \tag{5}$$

where K is a positive constant of proportionality.

Uncertainty is the average surprisal per token for an infinite length sequence of symbols. (It is always the receiver that is uncertain.)

Information (Entropy) By suitable choice of K, we may choose any base for the logarithm in Eq. 5. Here is the definition of entropy given by Shannon and Weaver [9]:

$$H(X) = -K \sum_{i=1}^{M} p_i \ln p_i. \tag{6}$$

Compare Eq. 6 to the equation for thermodynamic probability:

$$H(X) = -k \sum_{i=1}^{M} W_i \ln W_i, \tag{7}$$

where W_i is the thermodynamic probability of each state, k is Boltzmann's constant, equal to 1.3807×10^{-23} J/K, and H is the resultant entropy. The similarity between Eqs. 6 and 7 is striking.

Only absolute certainty banishes entropy absolutely In the event that there is total pattern redundancy in a communication, there is zero entropy. "For a given n, H is a minimum when all the P_i are epsilon [vanishingly small] but one. This is intuitively the most certain situation" [9].

The most uncertain situation has the maximum entropy "For a given n, H is a maximum and equal to $\log n$ when all the P_i are equal (i.e., $1/n$). This is also intuitively the most uncertain situation" [9].

Redundancy is the complement of entropy $H(X)$ related to its theoretical maximum, $\log N$:

$$R(X) = 1 - \frac{H(X)}{\log N}. \tag{8}$$

Redundancy $R(X)$ is what is left in a signal after subtracting its entropy. Information theory presents us with the somewhat counterintuitive outcome that the greatest amount of information is associated with the greatest degree of uncertainty. One way to view this is that entropy is the measure of the amount of information that is missing in the recipient prior to reception of the message.

While classical information theory is static, one-dimensional, and non-hierarchical, information theory offers crisp analogs to musical states of the listener: surprisal, expectation, and uncertainty. These theories help relate musical structure to the concomitant musical affect in the listener.

Conclusion I hope that these ideas can be used to help put music theory on an empirical basis. I believe that surprisal, expectation, and uncertainty are the universal underpinnings of music. I hope that this will encourage others to apply these ideas to the study of all forms of music.

References

1. Aristoxenus. C. 300 BCE: The Harmonic elements. In: Strunk, O. (ed.) Source Readings in Music History: Antiquity and the Middle Ages, pp. 27–31. W.W. Norton & Co., New York (1950)
2. Freyd, J.J.: Dynamic mental representations. Psychol. Rev. **94**(4), 427–438 (1987)
3. Freyd, J.J., Finke, R.A.: Representational momentum. J. Exp. Psychol. Learn. Mem. Cogn. **10**, 126–132 (1984)
4. Huron, D.: Sweet Anticipation: Music and the Psychology of Expectation. The MIT Press, Cambridge (2006)
5. Langer, S.: Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art. Harvard University Press, Cambridge (1957)
6. Meyer, L. B.: Emotion and Meaning in Music. University of Chicago Press, Chicago (1956)
7. Nyquist, H.: Certain topics in telegraph transmission theory. Trans. AIEE **47**, 617–644 (1928). Reprint as classic paper in: Proc. IEEE **90**(2), Feb 2002
8. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**(379–423), 623–656 (1948)
9. Shannon, C.E., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press, Urbana (1949)