Information Flow and Repetition in Music

David Temperley

Abstract A corpus analysis of common-practice themes shows that, when an intervallic pattern is repeated with one changed interval, the changed interval tends to be larger in the second instance of the pattern than in the first; the analysis also shows that the second instance of an intervallic pattern tends to contain more chromaticism than the first. An explanation is offered for these phenomena, using the theory of uniform information density. This theory states that communication is optimal when the density of information (the negative log of probability) maintains a consistent, moderate level. The repetition of a pattern of intervals is (in some circumstances, at least) highly probable; in some cases, the information density of such repetitions may be undesirably low. The composer can balance this low information by injecting a high-information (i.e., low-probability) element into the repetition such as a large interval or a chromatic note. A perceptual model is proposed, showing how the probabilities of intervals, scale degrees, and repetition might be calculated and combined.

Consider the melodies shown in Example 1. These five melodies all have several things in common. First and most obviously, each one involves a melodic pattern that is repeated in some way (marked with brackets above the score). In each case, the pattern is repeated at a different pitch level from the original, but the repetition maintains the same rhythm and, for the most part, the same pattern of generic intervals (i.e., intervals measured in steps on the staff). However, the pattern is not simply shifted along the underlying scale; in each case, it is slightly altered in some way. In the first melody, for example, a sixth in the first instance of the pattern (E♭4–C5) is changed to an octave in the second instance (F4–F5); similarly, in the third, fourth, and fifth melodies, an interval in the first pattern instance is replaced by a larger generic interval in the second instance. Other changes involve the addition of chromaticism: in the second and fifth melodies, diatonic scale degrees in the first instance of the pattern are replaced by chromatic degrees in the second instance.

There is another, more abstract commonality across these melodies that is perhaps less obvious than those mentioned above. Large intervals are

I am grateful to David Huron for making available to me his Humdrum encoding of the Barlow and Morgenstern corpus.

Journal of Music Theory 58:2, Fall 2014
DOI 10.1215/00222909-2781759

Example 1. (a) Mozart, "Non so piu" from *Le nozze di Figaro*, mm. 1–5; (b) Mozart, Piano Trio K. 548, II, mm. 1–2; (c) Beethoven, Sonata op. 2/3, I, mm. 1–4; (d) Schubert, Impromptu op. 90/4, mm. 47–50; (e) Brahms, Violin Sonata op. 100, I, mm. 1–10

less probable than small ones, and chromatic scale degrees are less probable than diatonic ones. (These claims are probably uncontroversial, but some evidence for them is given below.) Thus changing a small interval to a larger one is similar to changing a diatonic scale degree to a chromatic one: both kinds of changes *lower the probability* of the pattern. The thesis of this article is that this is a general characteristic of common-practice Western music: when an intervallic pattern is repeated with alterations, the alterations tend to lower the probability of the pattern rather than raising it.[1] (This is a tendency rather than a law; in some cases the opposite occurs. For example, the third instance of the pattern in Example 1d features a smaller interval than the previous one.) I will demonstrate that this tendency holds true statistically, and I will offer an explanation for it.

This study adopts a theoretical framework initially proposed in psycholinguistics, the theory of uniform information density (UID) (Levy and Jaeger 2007). In mathematical terms, the information carried by an element in a message is its negative log probability (see Figure 1). Informally speaking, the information of something represents how surprising it is. An element with a

1 This claim may also be valid for other musical styles, but I do not investigate that here.

Journal of Music Theory

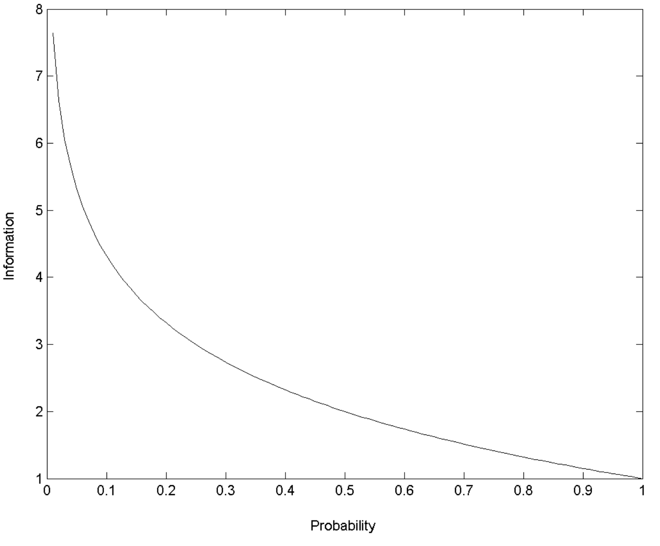


Figure 1. The relationship between probability and information

probability of 1 is completely predictable and thus carries no information; as elements decrease in probability, they become more surprising and hence more informative. Something with a probability of zero is completely surprising and thus conveys infinite information. (In technical terms, the information of an element corresponds to the number of bits—binary symbols—needed to represent it in the most efficient possible encoding of the entire language.)

Given a series of elements presented over time, we can define the *information density* (or *information flow*) as the amount of information per unit of time. Psycholinguistic studies have shown that low-probability elements (e.g., rare or unexpected words) take longer to process (Levy 2008); this suggests that there is an upper limit to the level of information density that human perceivers can easily absorb. But it is also, presumably, desirable for information flow to be fairly close to that limit, so as to maximize the amount of information conveyed. The UID theory thus states that it is preferable for information flow to maintain a fairly consistent, moderate level. One prediction that follows from this is that low-probability elements should be prolonged or spaced out in time more than high-probability elements. This prediction is confirmed by psycholinguistic studies showing that low-probability words and syllables in speech tend to be pronounced more slowly (Bell et al. 2003; Aylett and Turk 2004) and that optional words (e.g., *that* in the sentence “I knew that he was coming”) are more likely to be used when nearby words and syntactic structures are low in probability (Jaeger 2010). Similar phenomena have been observed in music as well: in a study of expressive performance,

Published by Duke University Press

Christopher Bartlette (2007) found that performers tend to take more time on unexpected (e.g., chromatic) harmonies.[2]

The use of information-theoretic and other probabilistic concepts in music research has a long history, going back more than half a century (Meyer [1957] 1967; Youngblood 1958), and has attracted renewed interest in recent years (e.g., Conklin and Witten 1995; Pearce and Wiggins 2006; Temperley 2007; Mavromatis 2009). Thus the application of UID to music would seem to be a natural topic for investigation. The focus of this study is not so much on the timing of events, as in previous UID research, but rather on a further prediction that follows from the UID theory—one that, to my knowledge, has not been widely explored. The prediction could be stated as follows: *When a message (or part of a message) is low in probability in one respect, it will tend to be high in probability in other respects, and vice versa.*

This prediction is quite general and could be interpreted musically in a variety of ways. To apply it to the current situation, we must distinguish between schematic probability and contextual probability. The *schematic probability* of an event is its probability in relation to the musical style as a whole; its *contextual probability* is its probability in relation to its specific context. I will argue that, in certain circumstances, it is highly likely for an intervallic pattern to be repeated: the contextual probability of melodic segments that repeat previous intervallic patterns is therefore quite high. The UID theory predicts that, to maintain an optimal level of information flow, the high contextual probability of these melodic segments should be balanced by a reduction in their schematic probability; this is accomplished by expanding their intervals and adding chromaticism.[3]

In what follows, I show, through a corpus analysis, that the relationships asserted above between repetition, interval size, and chromaticism do indeed hold true. I then offer an explanation for these phenomena, based on a model of melodic perception and the UID framework just described. The model is "tested" only in a qualitative, informal way; no rigorous evaluation of it is presented. Even so, it seems worthwhile to flesh out the model in some detail, as this raises some interesting and important issues in music perception.

2 One might wonder if constraints on *production* play a role here: perhaps speakers just need more time to pronounce uncommon words, and performers need more time to play unexpected harmonies. It seems unlikely, however, that the observed effects are due to such factors. In Bartlette 2007 the passages were simple tonal progressions, and the participants (graduate-level piano majors) were allowed to practice them as long as they wished before performing them; this suggests that fluctuations in timing were due not to difficulty but to an intentional expressive strategy. For discussion of this issue with regard to linguistic phenomena, see Jaeger 2010.

3 The distinction between contextual and schematic probability is similar to Eugene Narmour's (1990) distinction between "intra-opus" and "extra-opus" norms and to David Huron's (2006) distinction between "dynamic" and "schematic" expectations. The distinction is not clear-cut: one could say that even things such as interval size and scale degree, which I consider schematic, depend on context in a way (on the previous note and on the key, respectively). The term *contextual* is meant to imply a more specific prediction, unique to a particular context, such as the expectation for the repetition of a previous melodic pattern. One could also say that the role of contextual probabilities is itself schematic, that is to say, style dependent: motivic repetitions occur more often in some styles than in others.

Testing the predictions

I asserted in the second paragraph of this article that "when an intervallic pattern is repeated with alterations, the alterations tend to lower the probability of the pattern rather than raising it." We can express this idea more concretely in the form of two specific predictions. Think of each of the repeated patterns in Example 1 as a series of generic intervals; for instance, in Example 1d, the underlying pattern is +3 –1 (ascending fourth, descending step). The first prediction is that, when the second instance of the pattern changes one of the intervals, the changed interval in the second instance will tend to be generically larger than the corresponding interval in the first instance. The second prediction concerns chromaticism. We can think of any chromatic note in either the first or second instance of a repeated pattern as an alteration (assuming that the "default" version of the pattern stays entirely within the scale of the key). The prediction is that chromatic alterations will tend to occur in the second instance of the pattern more often than in the first. The objective is to test these two predictions.

This study uses a corpus of melodies taken from Harold Barlow and Sam Morgenstern's *Dictionary of Musical Themes* (1948), encoded on computer by David Huron in Humdrum format.[4] The corpus (hereafter the B&M corpus) contains 9,788 themes, all from instrumental pieces; most are opening themes, but the corpus also includes a significant number of other themes (e.g., second themes from sonata movements). The average length of themes is 4.9 measures; fewer than 2 percent of them are longer than 8 measures.[5] Table 1 shows the ten most frequently occurring composers in the corpus, with the number of themes by each one. To a first approximation, it seems reasonable to describe the corpus as a sample of "common-practice" themes, defining this term in the usual way to refer to eighteenth- and nineteenth-century Western art music. Fewer than 1 percent of the themes predate 1700. Roughly 20 percent of them date from after 1900; many of these are by conservative composers (e.g., Rachmaninoff and Respighi) whose style is rooted in the common practice.[6]

The first thing needed is a way of identifying repeated intervallic patterns. I initially define an "intervallic repetition" as any pattern of generic intervals that is immediately repeated, with exactly the same rhythm (this definition will be refined below). Intervallic repetitions are ubiquitous in

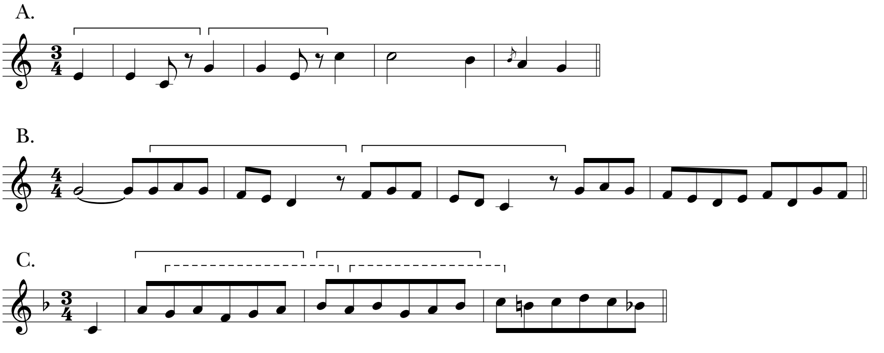
4 The dictionary can be accessed online at www.multimedialibrary.com/barlow.

5 The amount of music notated for each theme appears to be somewhat arbitrary and may be governed largely by the requirement that each entry should occupy exactly one line of notation. Some effort was apparently made to make entries correspond to musically meaningful units such as four-measure or eight-measure phrases, but the span of music included does not always correspond to what we would consider the "theme" (and this is somewhat subjective in any case). This does not appear to be a problem for this study.

6 As is discussed below, I exclude themes with changing time signatures and exotic rhythms such as quintuplets. This has the effect of excluding many of the themes by more progressive twentieth-century composers such as Bartók and Stravinsky (which is desirable, given that I want the corpus to represent common-practice music).

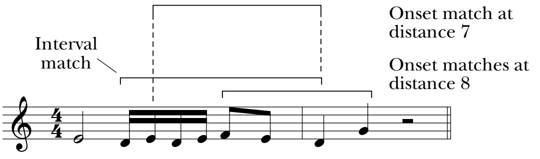
Table 1. The most commonly occurring composers in the Barlow and Morgenstern corpus

Composer	*Number of themes*
Mozart	587
Beethoven	568
Haydn	423
Brahms	383
Bach	375
Schubert	274
Handel	263
Schumann	238
Dvořák	210
Chopin	200



Example 2. (a) Haydn, Quartet op. 54/2, III, mm. 1–2; (b) Beethoven, Quartet op. 59/3, IV, mm. 1–4; (c) Beethoven, Sonata op. 2/1, III, mm. 41–43

common-practice music; diatonic sequences are perhaps the most obvious example, but they are often seen also in thematic passages that are not normally considered sequential. The openings of Mozart's fortieth, Beethoven's fifth, and Brahms's fourth symphonies are famous examples; Example 2 shows three others. Experimental work has shown that listeners can easily detect such repetitions—even when the pattern of *chromatic* intervals does not exactly repeat—and use them to encode melodies efficiently (Deutsch 1980). We define generic interval in the usual way, in terms of the number of steps on the staff; for example, the interval from F4 to A4 is a generic interval of +2 (an ascending third) and thus matches any other generic interval of +2, such as G4–B4, G4–B♭4, G♯4–B♭4, or G♯4–B♭♭4 (though the last of these is probably nonexistent). (The Humdrum encoding of the corpus includes pitch spelling information, allowing us to make enharmonic distinctions, e.g., distinguishing F4–A♭4, which is a third, from F4–G♯4, which is a second.)



Example 3. Beethoven, Sonata op. 2/3, I, mm. 1–2, showing onset matches at distances of 7 and 8. Only one of the three onset matches is an interval match: both notes are approached by a descending second. (The first note of the melody is not counted either in onset matches or in interval matches, as it is not approached by any interval.)

The concept of intervallic repetition will be further restricted in two ways. We may define the *distance* of a repetition as the temporal interval between corresponding notes in the two instances of the pattern (e.g., between the first note of the first instance and the first note of the second). While repetitions may occur at any distance, some distances are much more likely than others. In particular, it is especially likely for patterns to repeat at distances that correspond to levels of the metrical (and hypermetrical) structure. This is the case with the patterns in Examples 1 and 2, all of which repeat at distances of one or more measures. The general validity of this point can be demonstrated using the B&M corpus. Define an *onset match at a distance N* as a case where a note onset is followed by another note onset exactly *N* sixteenth-note beats later (see Example 3). Considering only melodies in 4/4 (with no metrical divisions below the sixteenth), the number of onset matches in the corpus was counted at each distance of one through thirty-two—that is, up to a distance of two measures (see Table 2). Then define an *interval match* as an onset match between a note *X* and a following note *Y* in which the generic interval to *X* is the same as the interval to *Y* (see Example 3). These data are also shown in Table 2. The probability of an interval match is highest at distances that correspond to levels of the 4/4 metrical grid—what we will call *metrically parallel* distances: 1, 2, 4, 8, 16, and (assuming a duple hypermetrical level) 32 sixteenths.[7] (This is true whether one considers the sheer number of interval matches or the proportion of onset matches that are interval matches, shown in the last column.) There is evidence also that listeners are particularly sensitive to metrically parallel repetitions. It has been observed that a melodic pattern that is performed in two

7 Distance 1 is a partial exception, as it is less frequent than several nonparallel distances in onset matches, though not in interval matches as a proportion of onset matches (IMs/OMs). The high values for distances 1 and 2 might be said to represent not so much pattern repetition as *inertia*, the tendency for an interval (especially a step) to be immediately followed by another interval of the same size and direction (Larson 2004). This process is a form of autocorrelation, a widely used technique in computational music research for problems such as meter identification (Brown 1993) and automatic transcription (Klapuri 2004).

Table 2. Onset matches (OMs) and interval matches (IMs) at different distances in themes in 4/4 in the B&M corpus

Distance (in 16ths)	*OMs*	*IMs*	*IMs/OMs*
1	**9,202**	**3,568**	**.388**
2	**23,085**	**7,650**	**.331**
3	7,166	1,852	.258
4	**31,542**	**9,365**	**.297**
5	6,399	1,397	.218
6	19,005	4,115	.217
7	6,248	1,392	.223
8	**30,332**	**9,103**	**.300**
9	5,817	1,253	.215
10	16,713	3,259	.195
11	5,128	985	.192
12	24,528	5,075	.207
13	4,791	957	.200
14	15,265	3,084	.202
15	4,964	1,108	.223
16	**25,856**	**8,448**	**.327**
17	4,486	1,081	.241
18	13,252	2,789	.210
19	3,654	725	.198
20	19,199	3,740	.195
21	3,121	649	.208
22	10,979	1,992	.181
23	3,011	592	.197
24	18,216	3,955	.217
25	2,580	537	.208
26	9,326	1,739	.186
27	2,215	425	.192
28	14,860	3,013	.203
29	1,907	419	.220
30	8,539	1,893	.222
31	2,153	538	.250
32	**16,527**	**6,459**	**.391**

Metrically parallel distances are shown in boldface.

different metrical contexts can seem like two totally different melodies (Povel and Essens 1985, 432; Sloboda 1985, 84); this is essentially what would occur if a pattern were repeated at a nonparallel distance (e.g., in a 4/4 context, a pattern that was repeated after seven sixteenths). I suspect that many nonparallel repetitions are not even perceived by listeners as repetitions.

For simplicity, I focus here on repetitions that occur at a distance of exactly one measure. While this excludes many significant repetitions, it has the advantage of being easily and objectively identifiable in all themes, regardless of the time signature. Themes that change time signature are a problem, as it is difficult to decide which distance to look at; such themes are disregarded in the counts reported below. (Themes with exotic time signatures, such as 5/4, and exotic rhythms, such as quintuplets, are also excluded, since they would greatly complicate the data collection process; some other themes are excluded due to errors or problems in the Humdrum encoding. Altogether, 627 themes are excluded for these reasons.)

While a pattern may repeat at a distance of one measure, this does not necessarily mean it begins on the downbeat; frequently it does not. The repeated pattern in Example 2a begins just before the downbeat; the one in Example 2b begins on the fifth eighth of the measure. To identify in an objective way where a pattern begins is a difficult matter. In Example 2c, the pattern could begin on the first downbeat (as shown by the solid brackets), but it could also begin on the second eighth of the measure (as shown by the dotted brackets).[8] We avoid this problem by considering only patterns that include all the notes of a single measure. Thus, in Example 2c, the intervallic repetition is the pattern marked by the solid brackets. Examples 2a and 2b contain no intervallic repetition; in Example 2a, for instance, the pattern of generic intervals within the first measure (–2 +4) is not repeated in the second measure (–2 +5). Again, this restriction causes many valid patterns to be missed (and may cause some dubious ones to be found), but it does not appear to introduce any *systematic* bias that would compromise the tests presented below.

To summarize: an intervallic repetition, by the current definition, consists of *a pattern of generic intervals, including all the notes of a single measure, that is exactly repeated in the following measure*. (Only patterns with at least two intervals—that is, three or more notes—are included; the reason for this will become clear below. The interval *to* the first note of each measure is disregarded, as this is not, strictly speaking, "within the measure.") I now turn to the first of the two claims above: when intervals are altered in a repeated pattern, they will tend to be larger in the second instance of the pattern than in the first. Define an *intervallic near-repetition* to be the same as an intervallic repetition except that exactly one of the generic intervals differs between the two instances, though it must be in the same direction in both cases.[9] Thus the

8 One could say that the identification of repeated patterns depends on the grouping structure (as defined by Lerdahl and Jackendoff 1983)—stipulating, perhaps, that a pattern must correspond to a low-level melodic "group." This seems doubtful in the case of Example 2c; the group of m. 1 surely includes the pickup note, but this is not part of the repeated pattern. In many cases the identification of groups is ambiguous and subjective. Grouping also depends partly on repeated patterns or "parallelism" (Lerdahl and Jackendoff 1983, 50–51); thus using grouping structure to identify repeated patterns creates a "chicken-and-egg" problem.

9 Thus a near-repetition must preserve the same note-to-note contour in both instances of the pattern. This restriction requires some explanation. Huron (2006, 75–77) notes that, in many musical idioms, the distributions for ascending and descending intervals are significantly different: skips tend to be ascending while steps are more often descending. This makes it problematic to compare ascending and descending intervals; for example, a large ascending skip may be more probable than a smaller descending skip. This asymmetry is only subtly evident in the B&M corpus (see Figure 3); however, it seemed safest to avoid the problem by comparing only intervals of the same direction. The requirement that an intervallic repetition (or near-repetition) must contain at least two intervals means that at least one interval will be matching between the two instances, even in a near-repetition. This is essential for the argument, since the changed interval is supposed to balance the low information of one or more repeated intervals.

themes in Examples 1b and 1d contain intervallic near-repetitions; Examples 1a, 1c, and 1e do not, because the repetition distance of the pattern is not one measure.[10] The claim is then that, in cases of intervallic near-repetition, the differing interval will tend to be larger (in generic interval size) in the second instance of the pattern than in the first. This was examined in the B&M corpus. I identified 676 cases of intervallic near-repetitions; in 391 of those cases (57.8 percent), the second interval was larger than the first. This differs significantly from the 50/50 distribution that would be expected if larger and smaller second intervals were equally common ($\chi^2(1) = 16.3$, $p < .0001$). A paired *t*-test across near-repetitions showed also that the differing interval was significantly larger in second instances than in first instances (mean generic size of differing interval size for first instances = 2.43, for second intervals = 2.72; $t(675) = -3.06$, $p < .005$).

The second claim made earlier was that the second instance of a repeated intervallic pattern is more likely than the first to contain chromatic notes—that is, notes outside the scale. The term *chromatic* is problematic in the case of minor mode, since the "scale" is ill-defined. To avoid this issue, we consider only melodies in major keys, where the distinction between diatonic (i.e., within-the-scale) and chromatic degrees is clear-cut. (The B&M corpus identifies the key of each theme—which is not necessarily the main key of the piece—allowing the scale degree of each note to be identified.) For simplicity, we classify scale degrees into twelve "neutral" categories, not distinguishing between, for example, ♯4 and ♭5.[11] Once again we test the prediction using the B&M corpus, this time examining intervallic repetitions rather than near-repetitions. Of the 1,046 intervallic repetitions in the corpus, in 122 cases the two instances of the pattern differ in the number of chromatic notes; of these, 74 have more chromatic notes in the second pattern instance (60.7 percent), significantly different from an even split ($\chi^2(1) = 5.12$, $p < .05$). A paired *t*-test on these cases shows that second pattern instances have significantly more chromatic notes than first instances (mean number of chromatic notes for first instances = 0.28, for second instances = 0.34; $t(121) = -2.79$, $p < .01$).

A more general way of framing the prediction about chromaticism would be to say that the second instance of a pattern will tend to contain less probable scale degrees than the first instance. This formulation of the prediction avoids an all-or-nothing distinction between diatonic and chromatic degrees, thus allowing the inclusion of themes in minor keys. The B&M cor-

10 In cases where a pattern occurs more than twice in succession, every pair of adjacent measures is counted; thus Example 1d contains a near-repetition between mm. 1 and 2, and another between mm. 2 and 3.

11 This seems like a reasonable simplification. Degrees within the scale are overwhelmingly likely to be spelled in the diatonic way; one is very unlikely to see degrees such as ♯7 (e.g., B♯ in the context of C major) or ♭4. Other neutral scale degrees may have multiple possible tonal interpretations (e.g., ♯1 vs. ♭2), but in that case both are chromatic. Even for the purpose of identifying the scale degree probability of notes, as I do below, representing them by neutral scale degrees should give similar results to representing them by tonal scale degrees; it may somewhat overstate the probability of chromatic notes, as it effectively sums together the probabilities of different tonal interpretations.

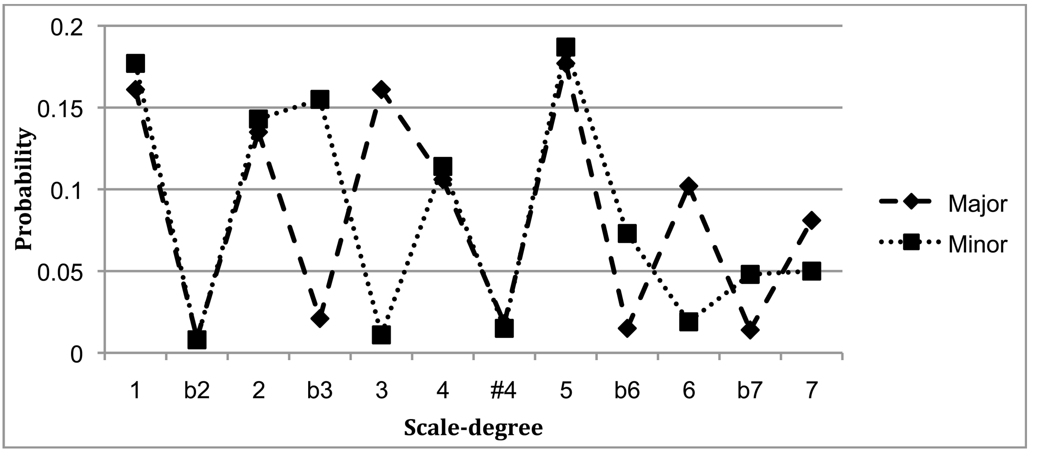


Figure 2. Key profiles from the Barlow and Morgenstern corpus, showing the distribution of scale degrees for major and minor keys

pus can be used to count the frequency of each scale degree, that is, pitch classes in relation to the tonic. Representing these frequencies as proportions of the total produces "key profiles" for major and minor keys, as shown in Figure 2; these are very similar to key profiles that have been proposed elsewhere, based on experimental work (Krumhansl 1990) and counts from other corpora (Temperley 2007). As expected, diatonic degrees have much higher probabilities than chromatic ones, and degrees of the tonic triad are somewhat more frequent than other diatonic degrees. The probability of a series of scale degrees can then be calculated as the product of their key-profile probabilities. To avoid the very small numbers that result from this, we take the logs of the probabilities; averaging these across the notes in a measure produces a *mean log scale degree probability* (MLSP) for each measure. The prediction is that the second pattern instance in an intervallic repetition will tend to have a lower MLSP than the first. (In this case we exclude cases in which the two pattern instances contain exactly the same sequence of scale degrees; in that case their MLSPs would be equal.) Testing the prediction on intervallic repetitions in the B&M corpus shows that the second instance has a lower MLSP in 496 of 879 cases, or 54.6 percent, significantly different from an even distribution ($\chi^2(1) = 14.28$, $p < .001$); a paired t-test shows that the MLSP of second instances is significantly lower than that of first instances (mean of first instances = −2.126, of second instances = −2.185; $t(878) = 3.76$, $p < .0001$). While this result undoubtedly reflects the higher level of chromaticism in second instances, it may also reflect distinctions between scale tones and, in particular, between tonic-triad degrees and other diatonic degrees. In Example 2c, while all notes of both pattern instances are within the scale, the second instance of the pattern has lower scale degree probability, probably because it contains only two tonic-triad notes whereas the first instance contains four. This is not a problem for the theory, which simply predicts that second pattern instances will have less probable scale degrees than the first, whether the distinction is between chromatic and diatonic degrees or between more or less probable diatonic degrees.

Before continuing, we should consider a possible alternative explanation for the results presented above. In general, it seems likely that the proportion of chromatic notes in a theme (defined as notes outside the scale of the main key of the theme) will tend to increase as it goes on. Some themes may actually modulate (though it appears that very few of the B&M themes do so within the portion included in the corpus); others may include applied chords and other sources of chromaticism that seem unlikely to occur at the very beginning of a theme. An investigation of this suggests that, indeed, scale degree probability does decrease slightly as themes continue; the average decrease in log scale degree probability from one measure to the next is .038 (this is just short of statistical significance: $r = -.56$, $p = .07$). To control for this, the test reported above was repeated, but with a difference: the mean MLSP for all measures in each position was calculated (the mean MLSP for m. 1, for m. 2, etc.), and, for each pattern instance, the mean for that measure position was subtracted from the MLSP.[12] The adjusted MLSP for second instances remained lower than for first instances, though the difference was somewhat reduced (first = −0.014, second = −0.041), and this remained significant ($t(877) = 1.68$, $p < .05$). Thus the general increase in chromaticism as themes continue appears to contribute to the difference between first and second pattern instances but does not entirely explain it. The same issue was investigated with interval size; in this case, the mean interval size does not generally increase as melodies go on (it actually *decreases* very slightly), so it seems clear that that does not explain the greater interval size of second versus first pattern instances.

A perceptual model

The task now is to explain the phenomena observed above. As stated in the first section, the argument hinges on the idea of uniform information flow: the high contextual probability of intervallic repetitions should be balanced by a reduction in their schematic probability, which is accomplished by expanding their intervals and adding chromaticism. But this leaves open the question of how schematic and contextual probabilities are calculated and com-

12 This was only done for mm. 1–8, since beyond that the counts of measures were considered too small for the mean MLSP values to be reliable; only intervallic repetitions entirely within that span of measures were included in the test.

bined by the listener in evaluating the overall probabilities of events. In what follows I propose a model of music perception that addresses this question. Two caveats: First, my intent here is not to model the precise quantitative results obtained in the last section but simply to show, in a general way, how the observed qualitative relationships among repetition, interval size, and chromaticism might be predicted. Second, the model is extremely oversimplified, in that it neglects many factors that undoubtedly affect melodic perception; despite this (as I show below), it is highly effective, both in modeling corpus data and in predicting perceptual judgments.

Let us begin with a simplified version of the model that considers just two factors: (1) generic melodic intervals and (2) scale degrees. It is well known that some intervals are more common than others in common-practice music and, in particular, that there is a preference for small intervals. Many theory textbooks (e.g., part-writing and counterpoint texts) express this principle explicitly, for example, recommending a preponderance of stepwise motion with only occasional leaps (Gauldin 1985, 17; Aldwell and Schachter 2003, 69). Corpus research has confirmed the preference for small and, in particular, stepwise intervals in a variety of musical idioms (Von Hippel 2000; Huron 2006). This principle has been found to influence perception as well, and in particular, melodic expectation, where it is often known as "pitch proximity"—the tendency for a melodic note to be followed by another one that is close in pitch. Virtually every model of melodic expectation incorporates this principle in some form (e.g., Narmour 1990; Schellenberg 1997; Lerdahl 2001; Larson 2004). The distribution of generic intervals over the entire B&M corpus is shown in Figure 3. With very few exceptions, the frequency of intervals decreases monotonically as they increase in size. The primary exceptions are

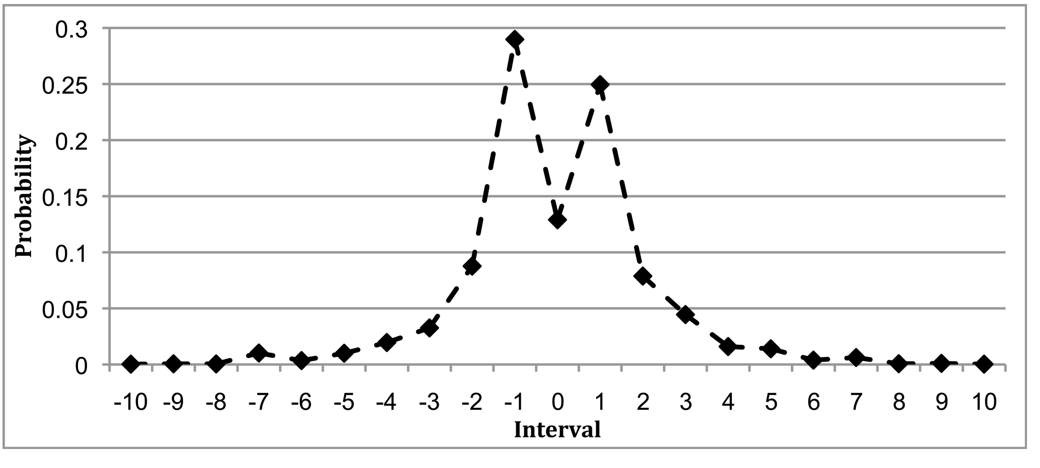


Figure 3. Distribution of generic melodic intervals in the Barlow and Morgenstern corpus

that the unison is somewhat less common than the ascending or descending step, and the octave is somewhat more common than the seventh, both ascending and descending. Apart from these caveats, the distribution can be modeled fairly closely with a normal distribution or "bell curve," and I will do so here.

Figure 3 identifies intervals only by their generic categories—not distinguishing, for example, between major and minor thirds. Introducing a second factor, scale degree, allows these finer distinctions. The idea that some scale degrees are more frequent than others is, again, uncontroversial, as both a fact about compositional practice and a factor in music perception. Carol Krumhansl's classic key profiles, based on experiments in which subjects judged the stability or "fit" of pitches in a tonal context, consistently reveal lower values for chromatic pitches; as Krumhansl (1990, 69–70) observes, listeners' internal key profiles most likely reflect the distribution of scale degrees in music they have heard. Most melodic expectation models also embody a preference for some scale degrees over others (Schellenberg 1997; Lerdahl 2001; Larson 2004). Scale degree distributions for the B&M corpus are presented in Figure 2; these will be used in the perceptual model as well.

What is needed now is a way of combining interval probabilities with scale degree probabilities to assign a probability to each note in a melody. In earlier work I have proposed a solution to this problem (Temperley 2007). Suppose we are in a C major context and the previous note in the melody is D4. The generic interval distribution (approximated with a normal distribution, as explained earlier) assigns an "interval probability" to each possible subsequent note; this is shown in Figure 4a for a two-octave range centered around D4. (It is an interesting question whether two specific intervals of the same generic category—for example, major and minor third—should be given the same probabilities at this stage, or different probabilities, reflecting their different chromatic sizes; I have chosen the latter course, as reflected in the figure.) We also take the appropriate scale degree distribution (the major one in this case) and duplicate it in each octave, as shown in Figure 4b. We then multiply the scale degree distribution with the interval distribution, normalizing the values of the combined distribution to sum to 1 (making it a well-defined probability distribution), as shown in Figure 4c. (The figure shows log probabilities rather than probabilities, since this more clearly represents distinctions between very small probability values.) In effect, the combined distribution gives highest probability to pitches that are both diatonic in the current key and close in pitch to the previous note—specifically, in the current case, C4, D4, and E4. Figure 4c also shows the actual distribution (log probabilities) of pitches in the B&M corpus for all major-key contexts following a pitch of scale degree 2. (The model treats all such contexts as the same.) Overall, the fit of the model to the data is quite close. Perhaps the most

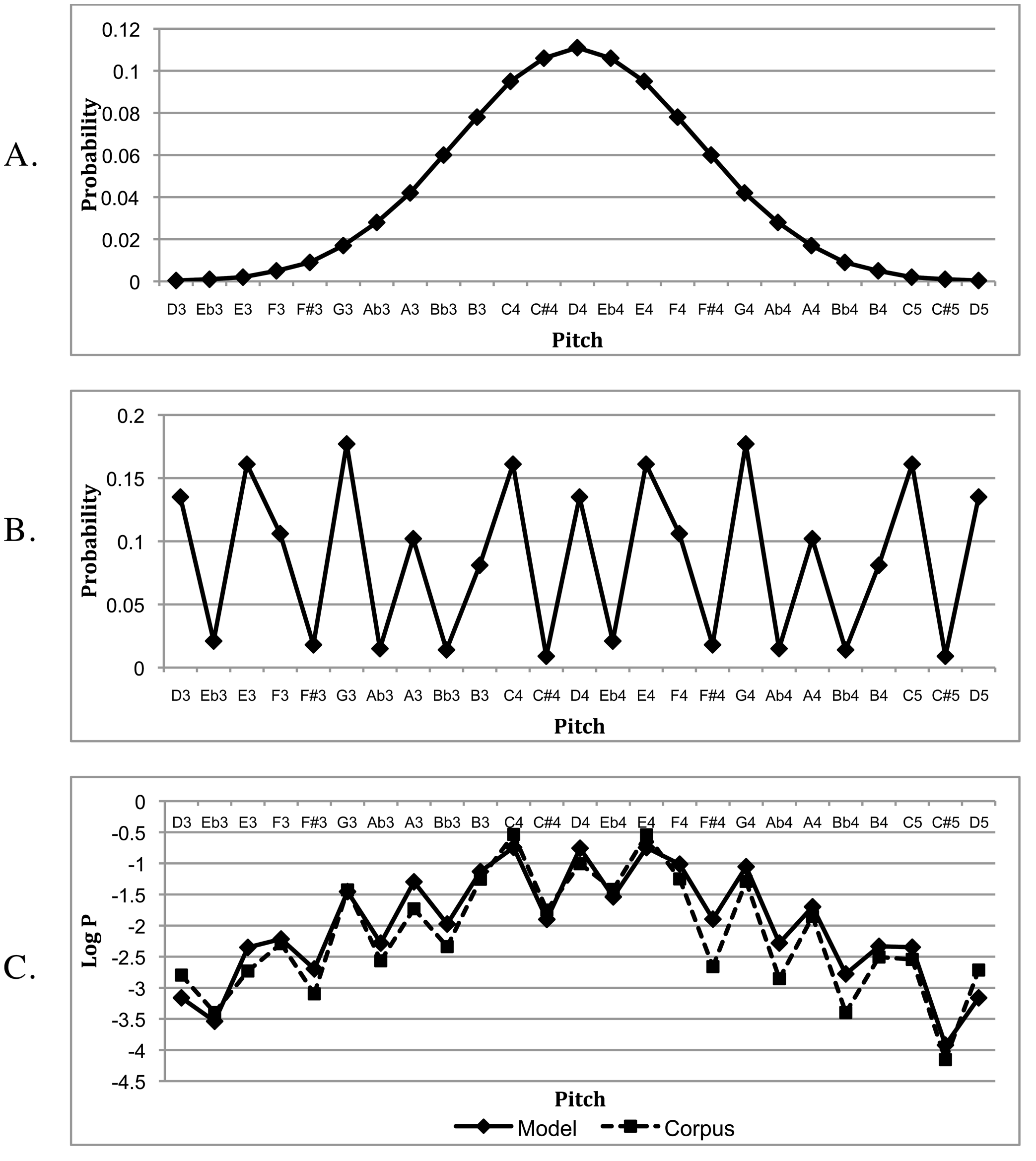


Figure 4. **(a) A distribution of chromatic intervals, assuming a previous pitch of D4. (b) The major-key scale degree distribution of the B&M corpus, assuming a key of C major, duplicated over two octaves. (c) The solid line ("model") shows the normalized product of a and b (log probabilities); the dotted line ("corpus") shows the distribution of pitches in major-key themes in the B&M corpus following scale degree 2 (assumed here to be D4 in C major).**

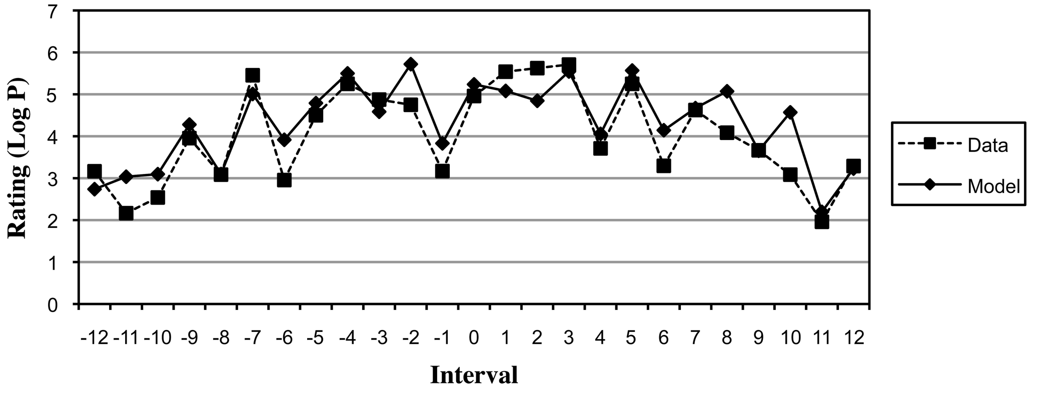


Figure 5. Data from a melodic expectation experiment (Cuddy and Lunney 1995) and a model's predictions. The data show judgments of expectedness (on a scale of 1 to 7) for pitches following a context of an ascending major second. Pitches are labeled in relation to the second pitch of the context; for example, –1 represents a descending minor second from the second pitch of the context. From Temperley 2007.

notable mismatch is the unison interval, whose probability is somewhat overestimated by the model.[13]

In Temperley 2007 I presented a model of melodic perception very similar to the one proposed here. (The main difference is that the earlier model also included a preference for each note to stay close to the center of the melody's overall range; this had a fairly small effect on the model's judgments.) The earlier model's predictions were compared with data taken from a melodic expectation experiment by Lola Cuddy and Carol Lunney (1995), in which listeners judged the expectedness of a note (on a scale of 1 to 7) following a two-note melodic context. (Since the key was not given in that case, the model had to guess the key from the context, just as listeners presumably did.) Experimental data (averaged over all subjects) and the model's predictions are shown in Figure 5 for a single context, an ascending major second. Over all contexts, the model yielded an excellent fit to listeners' judgments, with a correlation of .87. Thus the combination of interval size and scale degree seems to go a long way toward modeling both actual note probabilities (as reflected in compositional practice) and subjective probabilities (as reflected in expectation judgments).

We now turn to the problem of incorporating repetition into the model. We do this by modifying the interval distribution. The idea is that an interval that generically matches the interval at a previous metrically parallel position

13 The distribution of chromatic intervals across *all* contexts (not shown here) shows a roughly bell-shaped but uneven distribution; for example, there are many more major seconds than minor seconds. Similar results have been found in other corpora (Von Hippel 2000; Huron 2006, 74). Undoubtedly, this unevenness is due to the well-known fact that intervals differ in frequency within the diatonic scale; for example, there are five major seconds but only two minor seconds.

Example 4. A hypothetical melodic phrase with an indeterminate final note

should be given an increase in probability—I will call this the "repetition boost." One complication is that the interval distribution, as defined in Figure 4a, represents specific intervals, but the repetition boost depends on generic intervals: a third at the earlier position should raise the probability of any third at the later position. Here we simplify things by considering only the most common versions of each generic interval, that is, the ones that occur between notes of the diatonic scale: unison / m2 / M2 / m3 / M3 / P4 / A4, and their inverted and compound forms. Consider the melody in Example 4, and the probability distribution for the fourth note of the second measure. (Once again, we consider only parallelisms at a distance of one measure.) The previous note is scale degree 2; thus, without repetition, the interval distribution and combined distribution would be as shown in Figure 4, a and c. Given the repetition factor, the metrically parallel note in the previous measure, approached by an ascending third, would give rise to a repetition boost for every note in the interval distribution representing an ascending third—either m3 (F4) or M3 (F♯4), creating the distribution in Figure 6a. (A rough calculation—based on the values in Table 2—suggests that this boost should be about 0.1, but the exact value is not important here.) The adjusted interval distribution is then multiplied with the scale degree profile, creating the combined distribution shown in Figure 6b; the scale degree profile greatly reduces the value for F♯4 (as it is scale degree ♯4—not part of the scale), leaving F4 as the most likely note.

Imagine a series of notes that exactly repeated the generic intervals of the previous measure and remained entirely within the scale—for instance, Example 4 with F4 as the final note. In this case, every note would get the repetition boost, and all the note probabilities would be quite high. Recall that the UID theory states that information flow—information per unit time—should maintain a fairly constant rate, close to the optimal rate for human communication. While we do not know exactly what that optimal rate is, it seems reasonable to assume, if the theory has any validity, that the average information flow in musical themes is fairly close to it. One can imagine that, given the multiple repetition boosts, a melodic segment consisting entirely of repeated generic intervals and diatonic scale degrees might be somewhat below this optimal rate of information flow—or, in more informal terms, a bit dull and predictable. The composer might well decide to spice things up by changing one of the generic intervals; in this case, it would make sense to do

A.

B.

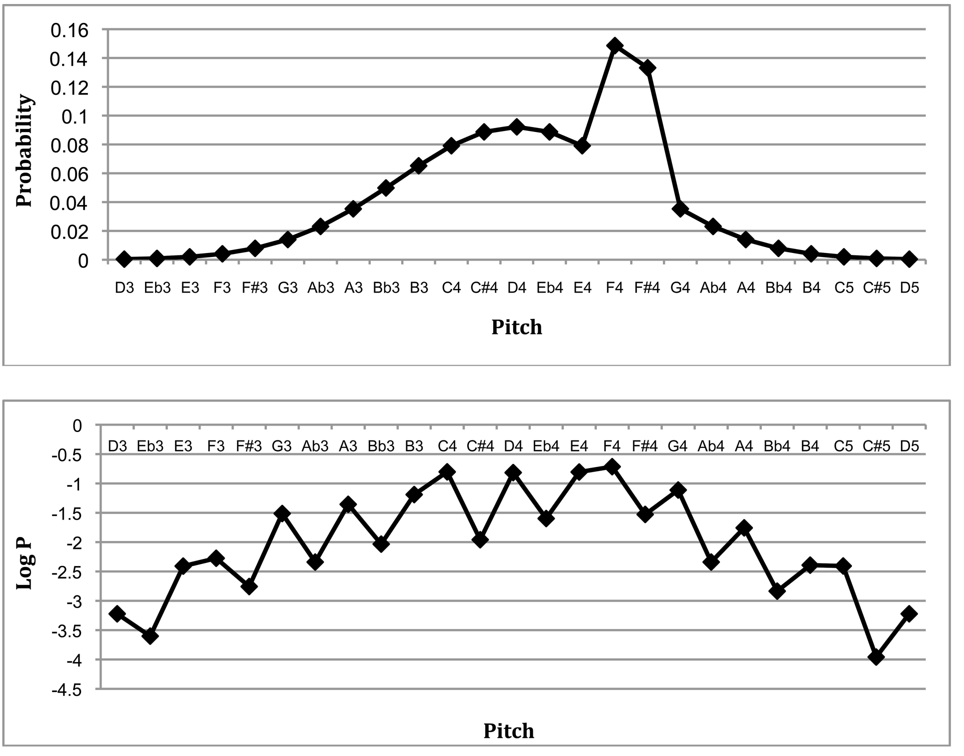


Figure 6. (a) A modified interval distribution, assuming a previous note of D4, including a "repetition boost" for ascending thirds; (b) the combined distribution resulting from the product of Figures 6a and 4b

so in a way that greatly increased the information content of the segment—that is to say, by choosing a relatively large interval (perhaps moving to A4 in Example 4 instead of the expected F4). Alternatively, the composer could maintain the same generic interval pattern but add chromatic inflections to one or more of the notes (perhaps moving to F♯ in Example 4 rather than F), thus increasing the information content in another way. In this way, the current model explains the tendency of common-practice composers to alter repeated patterns in ways that lower their schematic probability.

Discussion

In common-practice themes, in cases where a generic intervallic pattern is repeated with slight alterations from one measure to the next, there is a significant tendency for the second instance of the pattern to contain larger intervals and more chromaticism than the first instance. I have argued that this tendency can be explained in terms of the theory of UID. There is a high probability that the interval at one point in a measure will repeat the interval from the parallel point in the previous measure. Given a measure consisting entirely of such repeated intervals, the information flow might be undesirably low. This might well prompt the composer to alter the second instance of the pattern in some way that lowers its schematic probability, such as increasing the size of one of the intervals or adding a chromatic inflection to one or more of the notes. This provides an explanation for why the second instance of an intervallic pattern tends to have larger intervals and more chromaticism than the first.

I suggested at the outset that the claims being made about compositional practice are not laws but rather tendencies; the results presented above certainly bear this out. A great many themes in the corpus show patterns opposite to the ones predicted. With regard to interval size, in 285 of 676 near-repetitions, the changed interval is *smaller* in the second pattern instance than in the first, contrary to our prediction. A further point might be made as well: according to the current model (and disregarding chromaticism for the moment), it seems that an intervallic repetition will generally be more probable than any kind of near-repetition, even one in which the changed interval is smaller in the second instance than in the first. (This is due to the repetition boost; in Figure 6b, for example, the +m3 interval [going to F4] is more probable than the smaller +M2 [going to E4].) Thus, by the current reasoning, intervallic repetitions should be less desirable than near-repetitions, and therefore should be less common. But in fact, repetitions are *more* common than near-repetitions: there are 1,442 repetitions in the corpus but only 676 near-repetitions. Is this not a problem for the current argument?

Two points can be made in response. First of all, it is clear that the regulation of information flow is only one of the factors governing the construction of melodies; many other considerations also come into play. A melody must convey, or at least be compatible with, an underlying harmonic progression (I will return to this point). It may also participate in long-range linear processes and conventional schematic patterns. In Example 2c, it is difficult to think of an altered repetition of the first measure that would convey the underlying ii chord as clearly and elegantly as the unaltered repetition in the second measure. Similarly, the theme in Example 5a (containing a repeated one-measure pattern, and practically a two-measure pattern as well, except for the "missing" first note) conveys a 1–7–3–4 schema, a common strategy for opening themes (Gjerdingen 1987); any chromaticism or altered intervals might have hindered the identification of this schema. Apparently, such compositional imperatives sometimes—indeed, very often—override the preference for smooth information flow.

A second point, perhaps more interesting, is that there may be cases where an exact repetition is *preferable* from the point of view of information flow. I have assumed that the low information of some elements (i.e., notes) in a melody may be counterbalanced by the high information of other elements. But suppose we extend this reasoning to a higher level, defining "elements"

Example 5. (a) Beethoven, Quartet op. 131, IV, mm. 1–4; (b) Brahms, Piano Quintet op. 34, I, mm. 35–37; (c) Beethoven, Sonata op. 7, III, mm. 1–6

as larger melodic segments rather than as notes. It follows that, if a theme begins with a melodic segment that is particularly high in information content, UID would favor a following segment with very low information—perhaps an unaltered repetition of the first segment. Consider the second theme to the first movement of Brahms's Piano Quintet (Example 5b). The descending octave at the beginning of this motive is highly unusual and, no doubt, rather difficult to process; it is perhaps appropriate, then, that Brahms repeats the intervallic pattern in an unaltered form rather than altering it, as he often does—giving the motive a chance to sink in, one might say.[14]

If, as I have suggested, harmony can sometimes explain cases that go against the theory's predictions, it may also offer an alternative explanation for certain cases that the theory *does* predict. Return once more to Example 1c: I argued that the larger interval in the second pattern instance (a sixth instead of a fourth) can be explained as a means of increasing information content. But one might also note that repeating the fourth from the first pattern instance, yielding E4–A4, would conflict with the underlying I–V–V–I harmonic progression—a standard progression for the first four measures of a classical-period theme; the interval that is actually used, E4–C5, fits that progression. In this case, then, the altered repetition might well be explained in purely harmonic terms. In general, the assumption that melodies are constructed based purely on considerations of scale degree and interval size—as assumed by the model I present here—is, no doubt, a gross oversimplification of the compositional process. I would also suggest, however, that a view

14 The motive is slightly altered rhythmically: the first note is changed from an eighth note to a sixteenth note. In general, the phenomena posited here are particularly evident in Brahms's music: there are nearly as many near-repetitions, 66, as repetitions, 72, and among near-repetitions, the second instance contains the larger interval in 49 cases (74.2 percent).

of composition in which melodies are constructed to fit a predetermined harmonic progression is oversimplified as well. No doubt, both harmonic and melodic considerations operate simultaneously in the construction of themes, in a complex interactive process, and it seems reasonable to suggest that the shaping of information flow is one of the many considerations that affect this process.

While I have been focusing on harmony as a compositional factor independent of information flow, it may also affect information flow itself. It seems likely that the harmonic structure of a melody—whether explicit or implied—influences our expectations of what note will come next, thus affecting the informational contour of the melody. Consider Example 5c, containing a near-repetition between mm. 4 and 5. Measure 5 contains a smaller interval than m. 4; thus one could say that it increases the schematic probability of the pattern. But from a harmonic perspective, m. 5 could well be considered to be *lower* in probability than m. 4—in particular, because it contains a change of harmony on the second beat of the measure (from IV to I—confirmed by the accompaniment), the first time this has occurred in the piece. Similarly, in Example 1e, the higher information of the second five-measure phrase in relation to the first is due not only to its chromaticism but also to its dramatic harmonic "leaps"—from ii to IV/♭VII in mm. 6–7 and from a tonicized ♭VII to V7 in mm. 8–9. A more sophisticated model of melodic information flow would incorporate such harmonic considerations.

Quite apart from complicating factors such as harmony, this project could be extended in a number of ways. First of all, the UID theory points to several other possible predictions about the melodic parameters considered here—interval and scale degree. Analogous to the finding that low-probability words are pronounced more slowly, one might predict that notes approached by large intervals—being low in probability—would be longer than those following short intervals, and that chromatic notes would be longer than diatonic ones: that is, we would predict a positive correlation between interval size and duration, and a negative correlation between scale degree probability and duration. The problem with these predictions is that they are seriously confounded with other well-known musical principles. Since chromatic notes are generally ornamental in function, we might expect that they would normally be *short*. Indeed, the correlation between scale degree probability and length, across all notes in the corpus, is positive ($r = .31$), counter to the UID prediction. This does not rule out the possibility that UID exerts some pressure for chromatic notes to be long, but this pressure appears to be weak and outweighed by opposing forces. Possible confounds also arise with regard to interval size. As just noted, chromatic notes tend to be short; they are also (at least according to conventional theory) generally approached by stepwise motion. For this reason, a correlation between the length of a note and the size of the interval approaching it might arise indirectly. In this case, the alternative account makes the same prediction as the UID account: it predicts that notes approached by larger intervals will tend to be long. Across all notes of the corpus, there is indeed a positive correlation between length and interval size ($r = .69$), but it is difficult to know which of our two explanations is the real reason for it.

The treatment of melodic repetition in this study could also be further developed. Ideally, the "repetition boost"—the increase in probability for an interval that repeats a preceding one—would reflect not only repetition distances of one measure but other metrically parallel distances as well; at any given point, several generic intervals might be "boosted" due to their parallelism with intervals at different preceding points. One could also introduce distinctions between different pitch levels of repetition; it seems likely that repetition at the original pitch level (i.e., repetition of the pitches as well as the intervals) is more probable than at other levels, but this has not yet been systematically examined. It seems likely, also, that there is an "inertia" to repetition—that is, a single repeated interval increases the probability that the next one will be a repetition as well. (In Example 4, for instance, the fact that the first two intervals of the second measure repeat those of the first measure increases the probability that the third interval will also do so.) Such a factor would presumably increase the probability of multi-interval repetitions of the kind discussed here.

One could also expand the concept of altered repetition to include *rhythmic* alteration. The question of how to assess the probability of a rhythmic pattern is a difficult issue that I have explored elsewhere (see Temperley 2010). Apart from the probability of the rhythm itself, adding notes to a pattern will generally lower its probability. (There are many more eight-note patterns than four-note patterns, so there is less probability mass available for each one.) Thus the UID viewpoint predicts that a rhythmic alteration of a melodic pattern will usually add notes rather than remove them—an intuitively plausible prediction, in my view, but not yet tested.

From a broad perspective, this study highlights the complex, interactive relationship between composition and perception. Perception is shaped by composition, in that listeners' subjective probabilities and expectations are conditioned by the music they hear. But influence also flows the other way: composers are sensitive to listeners' capacities and preferences, and this in turn shapes *their* behavior. As a general statement about music composition, this may seem rather obvious. But to formulate it as a concrete, testable prediction is not such a simple matter, as I hope this study has shown. I have suggested that the theory of uniform information density offers a useful way of approaching this issue. UID provides a framework for formalizing the predictions made here regarding the relationships among interval size, chromaticism, and repetition; no doubt it could be applied to other musical questions as well. The UID theory also suggests that phenomena of this kind may not be unique to music but may connect in interesting ways with other domains (such as language) and with general principles of perception and cognition.

Works Cited

Aldwell, Edward, and Carl Schachter. 2003. *Harmony and Voice Leading*. Belmont, CA: Wadsworth Group/Thomson Learning.

Aylett, Matthew, and Alice Turk. 2004. "The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence, and Duration in Spontaneous Speech." *Language and Speech* 47/1: 31–56.

Barlow, Harold, and Sam Morgenstern. 1948. *A Dictionary of Musical Themes*. New York: Crown.

Bartlette, Christopher. 2007. "A Study of Harmonic Distance and Its Role in Musical Performance." Ph.D. diss., University of Rochester.

Bell, Alan, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Michelle Gregory, and Daniel Gildea. 2003. "Effects of Disfluencies, Predictability, and Utterance Position on Word Form Variation in English Conversation." *Journal of the Acoustical Society of America* 113/2: 1001–24.

Brown, Judith. 1993. "Determination of the Meter of Musical Scores by Autocorrelation." *Journal of the Acoustical Society of America* 94/4: 1953–57.

Conklin, Darrell, and Ian Witten. 1995. "Multiple Viewpoint Systems for Music Prediction." *Journal of New Music Research* 24/1: 51–73.

Cuddy, Lola, and Carol Lunney. 1995. "Expectancies Generated by Melodic Intervals: Perceptual Judgments of Melodic Continuity." *Perception and Psychophysics* 57/4: 451–62.

Deutsch, Diana. 1980. "The Processing of Structured and Unstructured Tonal Sequences." *Perception and Psychophysics* 28/5: 381–89.

Gauldin, Robert. 1985. *A Practical Approach to Sixteenth-Century Counterpoint*. Englewood Cliffs, NJ: Prentice-Hall.

Gjerdingen, Robert. 1987. *A Classic Turn of Phrase: Music and the Psychology of Convention*. Philadelphia: University of Pennsylvania Press.

Huron, David. 2006. *Sweet Anticipation*. Cambridge, MA: MIT Press.

Jaeger, T. Florian. 2010. "Redundancy and Reduction: Speakers Manage Syntactic Information Density." *Cognitive Psychology* 61/1: 23–62.

Klapuri, Aansi. 2004. "Automatic Music Transcription as We Know It Today." *Journal of New Music Research* 33/3: 269–82.

Krumhansl, Carol. 1990. *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Larson, Steve. 2004. "Musical Forces and Melodic Expectations: Comparing Computer Models and Experimental Results." *Music Perception* 21/4: 457–98.

Lerdahl, Fred. 2001. *Tonal Pitch Space*. New York: Oxford University Press.

Lerdahl, Fred, and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Levy, Roger. 2008. "Expectation-Based Syntactic Comprehension." *Cognition* 106/3: 1126–77.

Levy, Roger, and Florian Jaeger. 2007. "Speakers Optimize Information Density through Syntactic Reduction." In *Advances in Neural Information Processing Systems*, ed. Bernard Schölhopf, John Platt, and Thomas Hofmann, 849–56. Cambridge, MA: MIT Press.

Mavromatis, Panayotis. 2009. "Minimum Description Length Modeling of Musical Structure." *Journal of Mathematics and Music* 3/3: 117–36.

Meyer, Leonard. [1957] 1967. "Meaning in Music and Information Theory." In *Music, the Arts, and Ideas*, 5–21. Chicago: University of Chicago Press.

Narmour, Eugene. 1990. *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*. Chicago: University of Chicago Press.

Pearce, Marcus, and Geraint Wiggins. 2006. "Expectation in Melody: The Influence of Context and Learning." *Music Perception* 23/5: 377–405.

Povel, Dirk-Jan, and Peter Essens. 1985. "Perception of Temporal Patterns." *Music Perception* 2/4: 411–40.

Schellenberg, E. Glenn. 1997. “Simplifying the Implication-Realization Model of Melodic Expectancy.” *Music Perception* 14/3: 295–318.
Sloboda, John. 1985. *The Musical Mind*. Oxford: Clarendon.
Temperley, David. 2007. *Music and Probability*. Cambridge, MA: MIT Press.
———. 2010. “Modeling Common-Practice Rhythm.” *Music Perception* 27/5: 355–76.
Von Hippel, Paul. 2000. “Redefining Pitch Proximity: Tessitura and Mobility as Constraints on Melodic Intervals.” *Music Perception* 17/3: 315–27.
Youngblood, James. 1958. “Style as Information.” *Journal of Music Theory* 2/1: 24–35.

David Temperley is associate professor of music theory at Eastman School of Music.