Proceedings of the Annual Meeting of the Cognitive Science Society

Modeing Melodic Expectation: Using Three "Musical Forces" to Predict Melodic Continuations

Proceedings of the Annual Meeting of the Cognitive Science Society, 15(0)

Larson, Steve

1993

Modeling Melodic Expectation: Using Three 'Musical Forces' to Predict Melodic Continuations

Steve Larson
Center for Research on Concepts and Cognition
Indiana University
510 North Fess Street
Bloomington, IN 47408
(812) 855-6965
steve@cogsci.indiana.edu

Abstract

Part of what we call "expression" or "espressive meaning" in music may be regarded as an emergent property of the interaction of "musical forces" that I call gravity, magnetism, and inertia. These forces are implicit in Gestalt psychological principles of perceptual organization, current theories of tonal music, and recent experimental work in psychoacoustics. An explicit account of their operation and interaction allows us to predict which patterns of musical motion trained listeners will tend to expect in tonal music.

A computer program called What Next models the operation of these forces. Given a string of melodic pitches in a specific tonal context, What Next lists predicted continuations. A comparison of these predictions with the results of an experiment (Lake 1987), in which trained listeners were given a string of melodic pitches in a specific tonal context and asked to sing a continuation, suggests that the forces modeled have cognitive significance and explanatory power.

Music cognition

Music offers an enormously challenging area of study for the cognitive scientist. Passages of music—from the simplest melodic patterns to entire symphonies—give us a chance to listen to how the mind works.

One important aspect of listening to music is our tendency to anticipate what will happen next. Thus, music theorists have explored central questions about musical implications and realizations (Meyer 1956 and 1973; Narmour 1990 and 1992), music psychologists have gathered statistics on melodic expectancy (Carlsen 1988), and computer scientists have constructed models of tonal expectation (Bharucha and Todd 1989).

The purpose of this paper is to describe a theory of "musical forces" and to suggest how it may illuminate melodic expectation in tonal music. Elsewhere (Larson 1992), I have discussed the theory in greater detail, have shown how the results of psychological experiments make the operation of such "musical forces" seem plausible, and have suggested some implications for music teaching. In this paper, I describe a computer program (called What Next) that models aspects of this theory to make predictions about melodic continuations. These predictions may then be compared to the expectations of trained listeners as measured in a psychological experiment (Lake 1987).

Musical forces

This paper describes a theory about tonal music—that is, music of the western-European "concert" tradition from Bach to Brahms (music of the so-called "common-practice period") and much American popular music and jazz. Furthermore, the theory focuses primarily on melody (ignoring certain aspects of harmony and rhythm). The theory assumes that three "musical forces"—gravity, magnetism, and inertia—operate at all times on notes that are heard as unstable. Although there may be other musical forces, we can explain a great deal about tonal music with just these three.

Gravity is the tendency of an unstable note to descend to a *lower*, more stable pitch. For example, in a context where C is heard as stable and the D above it is heard as unstable, listeners experience musical gravity as a tendency of the D to descend to C.

Magnetism is the tendency of an unstable note to move (up or down) to the *nearest* stable pitch. Furthermore, magnetism is affected by distance—the closer we get to a goal, the more it attracts us. For example, in a context where D and G are heard as stable and the F between them is heard as unstable, listeners experience musical magnetism as a tendency of the F to ascend to G (because F is closer to G than it is to D). If that F should then move to F#, the magnetic force drawing us to G will intensify (because F# is now closer to G than F was).

Inertia is the tendency of a pattern of musical motion to continue in the *same* fashion. What is meant by "same" depends upon how that musical pattern is represented in our internal hearing. For example, if a pattern of musical motion begins "C-D-E, D-E-F", listeners may experience musical inertia as a tendency to continue the pattern "E-F-G", etc.

These definitions of gravity, magnetism, and inertia lead to two basic assertions. The first assertion is that melodic expectations in tonal music depend on the iterated operation of these forces on various hierarchical levels of musical structure.

By "iterated operation of these forces", I mean a multi-stage process like the following: (1) take a simple (but in some sense incomplete) melodic pattern, (2) follow the implications of one of the musical forces until a certain degree of stability is achieved, (3) take the resultant pattern, and (4) follow the implications of another of the musical forces until an even greater degree of stability is achieved.

By "hierarchical levels of musical structure", I mean the underlying pitch patterns of a melody. Some melodies may be viewed as embellishments of simpler patterns, which may themselves be viewed as embellishments, and so on (Schenker 1935/1979). This means that the musical forces may imply continuations of the simpler patterns or may imply continuations that further embellish melodies. Such a hierarchical view of musical structure is also necessary to interpret the meaning of "same" in the definition of inertia (which requires continuing patterns in the "same" way).

The second basic assertion is that goal-direction is a very important aspect of tonal music, and thus that the patterns of musical motion in which the final note is most strongly predicted by the musical forces are the most fundamental melodic patterns. Let's call the degree to which the forces predict the final note of a pattern the "strength" of a pattern. This implies that in highly goal-directed music, stronger patterns should occur more frequently than weaker patterns. As an example, consider the pattern G-F-G in the key of C, or more broadly, the scale degrees ^5-^4-^5 in any key (numbers preceded by a ^ refer to scale degrees). Because it defies all three forces, ^5-^4-^5 is quite weak: after ^5-^4, gravity predicts descent (to ^3); after ^5-^4, magnetism predicts motion to the closest stable pitch (to ^3); and after ^5-^4, inertia predicts motion in the same direction (to ^3). In fact, the pattern ^5-^4-^5 is less significant (and less common) in tonal melodies than the stronger pattern ^5-^4-^3, which gives in to all three forces. In fact, when ^5 is embellished with a lower neighbor, that lower neighbor is often raised chromatically so that the pattern is not ^5-^4-^5, but ^5-^#4-^5 (so that magnetism may be heard to overcome gravity and inertia). While a melody may go from ^4 to ^5 in the pattern ^3-^4-^5 (so that inertia may be heard to overcome gravity and magnetism), when a melody moves from ^5 to ^4, that melody more often continues by giving in to all three forces, as in the pattern ^5-^4-^3.

Many musicians—especially those with training in Schenkerian analysis (a method of discovering patterns at various hierarchical levels)—will find this musical explanation persuasive. But it has at least two drawbacks: it appears to involve circular reasoning and seems to confuse frequency with importance.

One reason that this musical explanation appears circular is that one cannot just pick some pieces and then count the frequency with which patterns occur in them. To count the number of patterns, one must first find them, a process that requires analytically separating the piece into patterns. Finding patterns on all hierarchical levels requires further analysis. In the end, such counting might prove more about the intellectual theory behind the analysis than about the aural and emotive experience of musical forces.

One reason that this musical explanation seems to confuse frequency with importance is that music does not always do what we expect it to do. It does not necessarily follow that just because a pattern is "more fundamental" it will happen more often. Music often creates some of its most salient effects by diverging from our expectations. In music, making clear distinctions between frequency, structural importance, and salience is a fascinating but complicated problem!

While this musical explanation may remain a persuasive one for many musicians, others may find more persuasive support for the theory in a comparison between the behavior of subjects in a psychological study and the behavior of a computer model based on the theory. This comparison exploits the importance of expectation, suggesting that stronger patterns play a more fundamental role in listeners' expectations.

A psychological study

In his dissertation study, William Lake (1987) asked music students at the University of Michigan to sing simple continuations. First, to establish a context, he played a chord and a scale for them. He then played a two-note beginning. Finally, he asked the students to sing that two-note beginning, "adding another tone or tones of your own choosing". After excluding the few responses in which subjects did not correctly reproduce the two-note beginning, he tabulated the frequency with which each third note (the first note of each continuation) was sung (excluding those notes sung in less than 12 per cent of the continuations).

On first inspection, Lake's results seem to support the theory advanced here—that is, the continuations seem fairly easy to rationalize in terms of musical forces. However, a computer model based on the operation of musical forces allows a clearer comparison. This comparison clarifies the operation of the musical forces, leads to the identification of a well-defined class of apparent exceptions, and suggests additional tests.

Figure 1: **Two displays of tonal pitch space (after Lerdahl 1988).**

The basic space, oriented to the tonic triad in C major.

C												(C)
C							G					(C)
C				E			G					(C)
C		D		E	F		G		A		B	(C)
C	Db	D	Eb	E	F	F#	G	Ab	A	Bb	B	(C)

A numerical representation of the basic space, oriented to the tonic major triad.

0												(0)
0							7					(0)
0				4			7					(0)
0		2		4	5		7		9		11	(0)
0	1	2	3	4	5	6	7	8	9	10	11	(0)

A computer model of musical forces

I have modeled the operation of these forces in a computer program called What Next. The following brief description of What Next not only explains the operation of the program but also clarifies aspects of the theory.

Given a string of melodic pitches, What Next lists predicted continuations. Melodic pitches are represented as integers (the tonic is represented as 0, and each other note is represented by an integer representing its distance in half steps above the tonic). Thus, for example, in the key of C, the stepwise pattern C-D-C-B-C would be represented as '(0 2 0 -1 0). This means that the program knows which key to hear the passage in, but that it does not distinguish between enharmonically-equivalent pitches such as G# and Ab.

Each prediction is generated by giving in to a specific musical force. "Giving in" to a force means moving (in the direction specified by the force) from an unstable pitch to a stable one. Thus, What Next must determine which notes are stable. In order to do this, the program uses a "tonal pitch space". Figure 1 shows the pitch space described by Fred Lerdahl (1988). This pitch space represents relative stability by means of the embedding of hierarchical levels. The pitches on any given level that are also contained in some "higher" (that is, more sparse) level are more stable than those not contained in that higher level. Thus, two levels of pitch space are required to specify the stability of a note: the level that contains that note (the "reference level") and the level that says "if that note is contained in this level too, then it is stable" (the "goal level").

The pitch space in What Next differs from Lerdahl's. His does not have an "obligatory register" (Schenker 1935). That is, it does not value the tonic of any single octave over that of another. (In fact, its use of modular arithmetic suggests that it is not a spatial model that possesses a real "up and down".) The notion of "obligatory register" suggests that if a piece sounds as though it should end on a particular C, then a C in a different octave will not provide as satisfactory an ending. What Next uses non-modular integers and an additional level (containing the tonic in a single octave) to represent the idea of "obligatory register".

Consider the two predictions illustrated in Figures 2 and 3. Both are predictions for what will follow the pattern G-A-? in the key of C.

Figure 2 shows a sample prediction based on the force of magnetism. Step one shows the motion G-A within the chosen reference level—the major scale. A goal level (that does not contain the second note of the pattern) is also chosen—the tonic triad. Step two shows the calculation of the distances (in half steps) to the closest stable pitches: G is two half steps away and thus closer than C, which is three half steps away. Thus, step three shows magnetism predicting motion within the reference level until the stable G is reached, resulting in the pattern G-A-G.

Figure 3 shows a sample prediction (for the same pattern, G-A-? in the key of C) based on the force of inertia. Step one shows the motion G-A within the same reference and goal levels—the major scale and tonic triad respectively. Step two shows inertia predicting motion will continue in the same direction, that is, upward by adjacent members of the reference level to B. But since B is not contained in the goal level, it is not stable. Therefore, inertia continues to operate on the growing pattern. Thus, step three shows yet another step, to C, resulting in the pattern G-A-B-C.

Figure 2: **Magnetism prediction for C: G-A-?**

Step One: The goal and reference levels are chosen.

```
goal level
(chord)
C             E         G                C

reference level
(scale)                  ----->
C      D      E  F      G      A      B  C
```

Step Two: The distances to the closest stable pitches are calculated (in half steps).

```
goal level
(chord)
C             E         G                C

reference level
(scale)                  <--2-   ---3--->
C      D      E  F      G      A      B  C
```

Step Three: The prediction is for motion (through the reference level) to the closest stable pitch (G)

```
goal level
(chord)
C             E         G                C

reference level
(scale)                  <-----
C      D      E  F      G      A      B  C
```

Resultant prediction: G-A-G.

Figure 3: **Inertia prediction for C: G-A-?**

Step One: The goal and reference levels are chosen.

```
goal level
(chord)
C             E         G                C

reference level
(scale)                  ----->
C      D      E  F      G      A      B  C
```

Step Two: Motion is continued in the same direction within the reference level, first to B (which is unstable).

```
goal level
(chord)
C             E         G                C

reference level
(scale)                         ----->
C      D      E  F      G      A      B  C
```

Step Three: Since B is also unstable, motion is again continued in the same direction within the reference level, now to C (which is stable).

```
goal level
(chord)
C             E         G                C

reference level
(scale)                               -->
C      D      E  F      G      A      B  C
```

Resultant prediction: G-A-B-C.

A sample of the program's output, for the same problem (G-A-?), is given in Figure 4. In addition to generating different continuations for a single beginning, the computer may also generate a given continuation by more than one method. The program groups its predictions in "trios". Each trio assumes its own combination of reference and goal levels, and each trio contains predictions based on the operation of the three different forces. (Remember that What Next represents pitches with integers: 7=G, 9=A, 11=B, and 12=C.)

In some cases, the program makes no inertia prediction. Since inertia is the tendency to continue *in the same way,* the program assumes that inertia develops only when notes move along in the *adjacent* pitches of a reference level. Thus, in the key of C, the notes C-D are adjacent in the reference level of the major scale, but not adjacent in the reference level of the chromatic scale. Therefore, they develop inertia to continue ascending in the major scale, but they develop no inertial tendency within the chromatic scale. The program indicates the absence of an inertia prediction by simply printing the pattern without a continuation.

Figure 4: **Sample output from What Next.**

```
>>> (what-next '(7 9))

The predictions for the pattern (7 9)
  at reference level major scale
      and goal level tonic octave
  are as follows--
    gravity prediction:  (7 9 7 5 4 2 0)
    magnetism prediction:  (7 9 11 12)
    inertia prediction:  (7 9 11 12)

  at reference level major scale
      and goal level tonic/dominant frame
  are as follows--
    gravity prediction:  (7 9 7)
    magnetism prediction:  (7 9 7)
    inertia prediction:  (7 9 11 12)

  at reference level major scale
      and goal level tonic triad
  are as follows--
    gravity prediction:  (7 9 7)
    magnetism prediction:  (7 9 7)
    inertia prediction:  (7 9 11 12)

Lake's results for the pattern (7 9) are
  (7 9 ((11 0.36) (7 0.31))).
```

In some situations, the operation of magnetism predicts two equally plausible continuations. This happens when the distance to the closest stable pitch above is the same as the distance to the closest stable pitch below. For example, in the key of C, where the reference level is the major scale and the goal level is the tonic triad, D can move up a whole step to E or down a whole step to C.

With each prediction, What Next prints Lake's statistics for comparison, but the program does not consult them when it generates a prediction.

A comparison of What Next with the results of Lake (1987)

Comparisons between the predictions of What Next and the continuations sung by Lake's subjects suggest that gravity, magnetism, and inertia play an important role in melodic expectation.

What Next applies the forces in only the simplest way. In applying gravity and magnetism it considers only the second note of the pattern. It does not perceive or create embellishments of hierarchical musical structure. And, in each prediction, it applies only one force only one time.

In light of these limitations, the performance of What Next is striking. Consider the seventy-five patterns that do not end on the tonic pitch. These patterns led subjects to sing up to four different third notes (excluding continuations sung less than twelve per cent of the time) for a total of 165 patterns. What Next predicts 132 of these (80%). The more that subjects agreed on a continuation, the more they agreed with What Next. For patterns in which subjects sang only one third note, the predictions of What Next included that continuation every time (100%). For patterns in which subjects sang two different third notes, What Next still did well (88%).

What Next did predict several continuations whose third notes appeared in less than twelve per cent of Lake's subjects' responses. The fact that many of these continuations seem musically plausible suggests that they might be useful in further tests of listener's judgments of melodic continuity.

It is also interesting to consider the third notes that appeared in more than twelve per cent of Lake's subjects' continuations but were not predicted by What Next. These seem to fall into two clear categories: those whose explanation seems to require the application of forces at another hierarchical level (especially those forming the "escape-tone pattern" up a half step then down three half steps—for example, E-F-D) and those that seem to require a different reference level not programmed into What Next (especially those that arpeggiate subdominant or dominant harmony, that is IV chords and V7 chords).

Since Lake's test used a major scale and chord to establish the context, it is not surprising that What Next seems to correspond best with Lake's subjects when it uses the major scale as its reference level and the major chord as its goal level (one type of exception here occurs when the two-note stimulus itself suggests a chord arpeggio, in which cases the best correspondences result from the use of the chord as reference level).

Finally, the theory seems to explain some results that may appear surprising. As Lake notes, the responses to two-note stimuli in which the second note functioned as ^4 seem to contradict rules of music theory requiring resolution of ^4 to ^3. But a consideration of the musical forces explains why ^4 has different melodic tendencies when preceded by different first notes.

Implications for future research

The success of What Next offers compelling support for the idea that musical expectations depend (at least in part) on the musical forces of gravity, magnetism, and inertia. It also suggests the value that computer models have for investigating music cognition.

What Next also raises some interesting questions. Can we quantify the relative impact of gravity, magnetism, and inertia? Would a computer model of the iterated application of musical forces at various hierarchical levels tell us more about musical forces and melodic continuations? What is the relation between key determination and melodic continuation? What, if any, aspects of the theory apply to music that does not fit the definition of "tonal" given above?

Acknowledgements

I would like to thank Bill Lake for providing me with a copy of his research results and for offering interesting and valuable advice.

Thanks also to Indiana University, whose Center for Research on Concepts and Cognition supported research on this project. My colleagues at CRCC, especially Doug Hofstadter, Gary McGraw, and James Marshall, offered thoughtful ideas on music and considerable help with computer programming.

What Next was written in MacScheme 4.0. Scheme is a dialect of LISP invented by Guy L. Steele, Jr. and Gerald Jay Sussman. Scheme for the Macintosh was developed at MIT (© 1984). MacScheme is published by Lightship Software, Inc. (© 1992).

References

Arnheim, Rudolph. 1974. *Art and Visual Perception: A Psychology of the Creative Eye (The New Edition).* Berkeley, Los Angeles, and London: University of California Press.

Arnheim, Rudolph. 1986. Perceptual Dynamics in Musical Expression. *New Essays on the Psychology of Art.* Berkeley, Los Angeles, and London: University of California Press.

Bharucha, Jamshed J. and Peter M. Todd. 1989. Modeling the Perception of Tonal Structure with Neural Nets. *Computer Music Journal* 13/4. Reprinted in Todd and Loy 1991.

Carlsen, James C. (editor). 1988. *Melodic Expectancy: A Special Issue of Psychomusicology.*

Lake, William. 1987. Melodic Perception an Cognition: The Influence of Tonality. PhD dissertation, University of Michigan.

Larson, Steve. 1992. Scale-Degree Function: Cognition Research and Its Application to Aural-Skills Pedagogy. Presented to the 1992 annual meeting of the Society for Music Theory and available (from the Center for Research on Concepts and Cognition; Indiana University; 510 North Fess Street; Bloomington, IN 47408) as CRCC Technical Report #67.

Lerdahl, Fred. 1988. Tonal Pitch Space. *Music Perception* 5/3: 315-349.

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

Meyer, Leonard. 1956. *Emotion and Meaning in Music.* Chicago: University of Chicago Press.

Meyer, Leonard. 1973. *Explaining Music.* Berkeley: University of California Press.

Narmour, Eugene. 1990. *The Analysis and Cognition of Basic Melodic Structures.* The University of Chicago Press.

Narmour, Eugene. 1992. *The Analysis and Cognition of Melodic Complexity.* The University of Chicago Press.

Schenker, Heinrich. 1935/1979. *Free Composition: Volume III of New Musical Theories and Fantasies (Der freie Satz).* Translated and edited by Ernst Oster. New York and London: Longman.

Todd, Peter M. and D. Gareth Loy (editors). 1991. *Music and Connectionism.* Cambridge, MA and London: The MIT Press.