Redefining Pitch Proximity: Tessitura and Mobility as Constraints on Melodic Intervals

Author(s): Paul Von Hippel


Published by: University of California Press


Accessed: 11-06-2015 01:29 UTC
Redefining Pitch Proximity: Tessitura and Mobility as Constraints on Melodic Intervals

PAUL VON HIPPEL
Ohio State University

In descriptions of melodic structure, *pitch proximity* is usually defined as the tendency for small pitch intervals to outnumber large ones. This definition is valid as far as it goes; however, an alternative definition is preferable. The alternative defines pitch proximity in terms of two more basic constraints—a constraint on *tessitura* (or pitch distribution) and a constraint on *mobility* (or freedom of motion). This new definition offers several advantages. Whereas the usual definition predicts only interval size, the new definition predicts interval direction as well. The usual definition predicts small intervals generally, whereas the new definition predicts context-sensitive variations in interval size. Finally, if the new definition is given the first few notes in a melody, it can assign a probability to each of the pitches that could occur next. In sum, the new definition offers a more precise and detailed description of melodic structure.

Received February 22, 1999; accepted for publication September 29, 1999.

Perhaps the simplest property of melodic structure is the rule known as *pitch proximity*. This rule has traditionally been defined as a tendency for small pitch intervals to outnumber large ones. A preponderance of small intervals has been measured in a wide variety of musical cultures, including indigenous melodies from Europe, Africa, North America, and the Caribbean (Dowling, 1967; Ortmann, 1926; Watt, 1924; Zipf, 1949).1 A classic demonstration by Watt (1924) is displayed in Figure 1.

Although the traditional definition is valid as far as it goes, I will argue that an alternative definition is preferable. This alternative defines pitch proximity as the result of two more basic constraints. The first constraint affects a melody’s distribution of pitch heights—in a word, its *tessitura*.

1. Dowling’s demonstration relied on data reported by Fucks (1962) and Merriam (1964). The data reported by Merriam, in turn, was originally collected by Merriam, Whinery and Fred (1956).

Address correspondence to Paul von Hippel, School of Music, Ohio State University, 1866 College Road, Columbus, OH 43210. (e-mail: von-hippel.1@ohio-state.edu)
Paul von Hippel

Fig. 1. In an analysis of 56 Lieder by Schubert, Watt (1924) found that small melodic in-

tervals greatly outnumber large ones. Similar results, less fully tabulated, were obtained for the
songs of two Native American tribes—the Ojibway and the Lakota.

The second constraint affects a melody’s freedom of movement within its
tessitura—in a word, its mobility. These two constraints, on tessitura and
mobility, limit a melody’s choice of intervals.

At first glance, the proposed change of definition may appear trivial. I
will demonstrate, however, that the new definition offers important ad-

avantages over the old one. Whereas the old definition can predict only the sizes
of melodic intervals, the new definition predicts their directions as well.
The old definition predicts small intervals regardless of context, whereas
the new definition predicts systematic variations in interval size. Finally,
the new definition fits a statistical formalism that, given the first several
notes in a melody, assigns a probability to each of the pitches that could
occur next. This formalism and its corollaries will be developed in the pages
that follow.

**Constraint 1: Tessitura**

To begin, we can observe informally that every melody is limited in its
range of pitches, and that most melodies seem to favor the center of their
range. This preponderance of moderate pitch heights could have a variety
of causes. On many instruments, pitches of moderate height are easier to
play. Even when extreme pitches can be played easily, they tend to produce
less definite pitch perceptions (Terhardt, Stoll, & Seewann, 1982). Moreover,
because melodies move primarily by small interval, melodies must
generally pass through the center of their range in order to reach the ex-
tremes. Whatever the cause, there does seem to be a tendency for melodies to use the fringes of their range sparingly. It is for this reason that musicians use the word "tessitura," which "differs from range in that it does not take into account a few isolated pitches of extraordinarily high or low pitch" (Apel, 1969, p. 839).

To make these observations more concrete, tests were conducted on a diverse sample comprising 176 indigenous folk songs from four different continents: 30 songs from a Chinese collection (Chung-kuo yin yüeh yen chiu so [Chinese Music Research Institute], 1959), 24 songs from a South African collection (Makeba, 1971), 42 songs chosen at random from a Native American Ojibway collection (Densmore, 1910, 1913; computer data files available in von Hippel, 1998), and 80 from a European collection (Schaffrath, 1995)—4 songs chosen at random from each of 20 regions of Europe. One of the European folk songs is notated in Figure 2, and a tally of the song's pitch heights is displayed as a bar graph in Figure 3. The bar graph's mounded shape shows clearly that most of this melody's pitches occur toward the middle of its range. More precisely, the pitch heights in this melody—like those in the other sampled melodies—approximate a normal or Gaussian distribution.2

It is evident from Figure 3 that the size of melodic intervals is constrained by the pitch distribution. Given the pitches used in this melody, the largest possible interval is 14 semitones—and that interval can be formed only by straddling two rare pitches at opposite extremes of the range. In contrast, there are many ways to form smaller intervals using more abundant pitches near the middle of the range. To be more precise, we might point out that, given a normal distribution of pitch heights, about two thirds of a melody's

![Fig. 2. A German folk song ("Lasset uns schlingen dem Frühling Blümlein") from the Essen Folksong Collection (Schaffrath, 1995).](image)

---

2. As a simple test of normality, we can calculate skew and kurtosis values for each of the 176 sampled melodies. Values of 0 would be consistent with a normal distribution, whereas a consistent tendency toward positive or negative values would indicate a departure from normality. According to a sign test, the skew of these melodies did not tend significantly toward positive or negative values (75 positive, 100 negative; \(\chi^2 = 3.2, p = .07\)). The kurtoses did exhibit a significant tendency toward negative values (23 positive, 153 negative; \(\chi^2 = 95, p < .0001\)), but the median kurtosis (−.7) did not seem large enough to warrant adjustments.
pitches will be within a standard deviation of the mean pitch. The standard deviation of a melody's pitch heights—3.5 semitones for the melody in Figure 2—can therefore serve as a useful index of tessitura. Median values for the tessitura index are 3.1 semitones for the European folk songs, 3.8 for the Chinese, 4.0 for the South African, and 4.7 for the Ojibway. (The median ranges are roughly four times as large: 12 semitones for the European folk songs, 15 for the Chinese, 17 for the South African, and 16.5 for the Ojibway.)

The effect of tessitura on interval size can be measured by transforming each of the sampled melodies into a "scrambled twin." A scrambled twin is produced by simply reordering a melody's pitches at random. The resulting twin is identical to the original melody in its pitch distribution, but random in its pitch order—random, that is, in its intervals. The intervals present in a scrambled twin are the exclusive result of the pitch distribution—a result, that is, of the scale, the relative prevalence of different scale degrees, and the tessitura. If interval size is constrained by tessitura, we would expect the scrambled twins to exhibit a preponderance of small intervals.

For the purpose of this comparison, an interval can be defined as the semitone distance between any consecutive pitches that are not separated by a phrase boundary or rest. To preserve the number of intervals present in the original melodies, scrambled twins can be constructed in such a way that the positions of rests and phrase boundaries are fixed, even as the positions of pitches are randomized. Scrambled twins of this type were constructed for each of the 176 sampled melodies. Tallies of noncompound
Redefining Pitch Proximity

Fig. 4. Noncompound interval sizes in folk songs from four different continents are compared with interval sizes in corresponding sets of "scrambled twins." Each scrambled twin was composed by rearranging the pitches of the corresponding folk song in a random order. The prevalence of small intervals in the scrambled twins can result only from constraints on tessitura. The greater prevalence of small intervals in the original folk songs can be attributed to constraints on mobility.

intervals in the original melodies and their scrambled twins are displayed in Figure 4.

To begin by examining specific intervals, the distributions for the scrambled twins display sharp peaks at the consonant intervals of 5 and 7 semitones (the perfect fourth and perfect fifth), as well as a deep trough at the dissonant interval of 1 semitone (the minor second). These peaks and troughs are probably an artifact of scale structure. Both the pitch content of common scales and the relative prevalence of certain scale degrees tend to maximize the possibility of consonant intervals and minimize the possibility of dissonant intervals (Huron, 1994; Smith, 1997).

Setting aside these detailed observations, Figure 4 exhibits a coarser general pattern: In all four sets of scrambled twins, small intervals greatly outnumber large ones. This pattern is consistent with the simple claim that the constraint on tessitura, by itself, would tend to produce an excess of small intervals.
Constraint 2: Mobility

Although small intervals are prevalent in the scrambled twins, they are considerably more prevalent in the original melodies. The comparison is summarized in Table 1, which shows that the mean interval size for each repertoire is one to four semitones smaller than the mean interval size for that repertoire’s scrambled twins. When the 176 original melodies are considered individually, all but 5 of them have a smaller mean interval size than the corresponding scrambled twin. These results show that, when the pitch distribution is controlled, melodic intervals remain considerably smaller than would be expected by chance.

This simple analysis corroborates the intuition that interval size is subject to a second constraint beyond that of tessitura. This second constraint limits the freedom with which a melody can move through its tessitura—it is, in other words, a constraint on mobility. The effect of mobility constraints can be described by the observation that, even allowing for the compact distribution of a melody’s pitch heights, each pitch tends to be close to the pitch before. More formally, we would expect most melodies to have a positive correlation between the heights of consecutive pitches—that is, a positive lag-one autocorrelation. The strength of this autocorrelation may serve as an index of constraints on mobility.

For all but one of the 176 melodies in our sample, the lag-one autocorrelation between consecutive pitch heights is positive. When pitch height is measured in semitones, the median value of this autocorrelation is .46 for the Chinese folk songs, .59 for the European folk songs, .59 for the South African folk songs, and .84 for the Ojibway folk songs. Notice that the Ojibway folk songs have the most constrained mobility; this explains how they can have the widest tessitura of the four repertoires (median index, 4.7 semitones), yet still have the smallest mean interval size (1.6 semitones).

<table>
<thead>
<tr>
<th>Folk Song Repertoire</th>
<th>Mean Interval Sizes (Semitones)</th>
<th>Melodies With Smaller Mean Interval Sizes Than Their Scrambled Twins</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>in Original Melodies in Scrambled Twins</td>
<td></td>
</tr>
<tr>
<td>Chinese</td>
<td>2.9 3.5</td>
<td>26 out of 30</td>
</tr>
<tr>
<td>European</td>
<td>2.2 3.7</td>
<td>79 out of 80</td>
</tr>
<tr>
<td>South African</td>
<td>2.3 4.3</td>
<td>24 out of 24</td>
</tr>
<tr>
<td>Ojibway</td>
<td>1.6 5.5</td>
<td>42 out of 42</td>
</tr>
</tbody>
</table>

Table 1
Mean Interval Sizes in Four Folk Song Repertoires and in Their Scrambled Twins
To illustrate this index of mobility, consider again the melody displayed in Figure 1. In Figure 5, we have plotted each pitch in this melody against the pitch that succeeds it. The strong lag-one autocorrelation between successive pitch heights (.70) indicates that, even compared with the rest of the European sample, the mobility of this folk song is highly constrained.

Forecasting Pitch Heights

Tests of the four sampled repertoires indicate that the autocorrelation between successive pitch heights is approximately linear. That is, on a scatterplot like Figure 5, the points fit a straight line rather than a higher order curve. The existence of such a linear autocorrelation suggests that the height of any pitch can be predicted, with some error, by a first-order autoregression formula. In particular, if the first several notes in a melody \((P_1, ..., P_{i-1})\) have a mean of \(m\) semitones and a lag-one autocorrelation of \(r_1\), the most likely height for the next pitch \(P_i\) will fit the following equation:

\[
P_i - m = r_1 (P_{i-1} - m)
\]

(1)

where, in practice, the predicted pitch height \(P_i\) would be rounded to the nearest discrete pitch.

Put into words, this equation may be interpreted in the following terms. After a high pitch (i.e., when \(P_{i-1} - m > 0\)), the next pitch tends to be somewhat lower (i.e., closer to the mean by a factor of \(r_1\)). Similarly, after a low pitch, the next pitch tends to be somewhat higher. Both of these tendencies are stronger when the antecedent pitch is extreme (i.e., \(P_{i-1}\) is distant from \(m\)) or when the constraint on mobility is weak (i.e., \(r_1\) is small).

These phenomena, familiar to statisticians as “regression toward the mean,” are straightforward results of the constraint on tessitura. Simply put, after an extreme pitch—or, for that matter, before it—most of the available pitches are quite a bit closer to the middle of the tessitura. For this reason, upward intervals, especially large ones, tend to begin low and end high, whereas downward intervals, especially large ones, tend to begin high and end low. These tendencies are illustrated in the two panels of Figure 6, which shows that for the German folk song in Figure 1—as for all of the 176 sampled melodies—intervals (as measured in semitones) exhibit

3. As an informal test of linearity (Darlington, 1990), we examined residual scatterplots for eight randomly selected melodies (two from each of the four repertoires); these scatterplots exhibited no obvious signs of curvilinearity. As a more formal test (Darlington, 1990), we tried to fit quadratic curves to the scatterplots for each of the 176 melodies. In general, these curves explained only marginally more of the variance than could be explained by a straight line (median increment in \(R^2 = .01\)). It therefore seems that the correlation between successive pitch heights could be reasonably characterized as linear.
a negative correlation with the height of their starting pitch, and a positive correlation with the height of their ending pitch.

The tendencies illustrated in Figure 6 have a number of straightforward corollaries. One corollary is that the peak pitch in a phrase—which is, by definition, high—tends to be approached and left by skip. A second corollary is that after a skip—which tends to land on an extreme pitch—a melody is likely to retreat by changing direction. Although both of these tendencies have been claimed as independent principles of melodic structure (Eitan, 1993, 1997; Meyer, 1956; Nanino & Nanino, ca. 1600; Narmour, 1990; Toch, 1948), recent analyses suggest that they are mere side effects of constraints on tessitura (Huron, 1999; von Hippel & Huron, 1999).

Assigning Probabilities to Pitches

A test of the scatterplot in Figure 5—and similar tests for the other sampled melodies—indicates that the variances around the autoregression line are
Fig. 6. The pitches and intervals plotted here come from the melody in Figure 3; each point’s area is proportional to the number of times that the corresponding pitch-interval succession occurs in the melody. Because of constraints on tessitura, the melody’s larger intervals tend to be associated with relatively extreme pitches. In particular, large upward intervals tend to begin on low pitches (left) and/or end on high pitches (right). Similarly, large downward intervals tend to begin on high pitches (left) and/or end on low pitches (right). (Note. In keeping with this paper’s definition of interval, this plot does not include intervals that cross phrase boundaries or rests. As a result, this plot contains fewer data points than Figure 5, which plotted every succession of pitches, regardless of what comes between them.)
roughly homogeneous. That is, regardless of the antecedent pitch height, the heights of consequent pitches are about equally variable. Given this property, our autoregression formula can do more than simply predict the pitch that is most likely to occur next—the formula can also assign a probability to the subsequent occurrence of any given pitch. In particular, if the first several notes of a melody ($P_1, \ldots, P_{i-1}$) fit a normal distribution with mean of $m$, a standard deviation of $s$, and an autocorrelation of $r_1$, the probabilities for the next note will fit a normal distribution. The mean of this probability distribution, as indicated above in equation (1), will be $P_i = r_1(P_{i-1} - m)$ semitones, and the standard deviation will be $\sqrt{1-r_1^2}$ semitones (Brockwell & Davis, 1996). Note that the width of the probability distribution ($\sqrt{1-r_1^2}$) depends on both tessitura and mobility; in particular, pitches are predicted with greater certainty when the tessitura is narrow ($s$ is small) or the mobility is highly constrained ($r_1$ is large).

In practice, this continuous probability distribution will be broken into discrete bins corresponding to the available pitch classes. An example is displayed in Figure 7. Here the melody from Figure 2 has been interrupted at the end of its sixth measure, and a probability distribution has been generated for the following pitch. Because the mobility is constrained, the peak of the probability distribution is near the most recent pitch; moreover, because the most recent pitch is high in the tessitura, the peak of the probability distribution is somewhat lower.

For melodies with a downward trend—for example, most Ojibway melodies (Densmore, 1913)—the next note will usually be lower than our formula predicts. But for a melody that maintains a consistent tessitura, the formula's predictions should be roughly correct—at least with respect to pitch height. In Figure 7, for example, it is encouraging to observe that the

![Fig. 7. Here the melody from Figure 2 has been interrupted at the end of its sixth measure, and our definition of pitch proximity has been used to generate a probability distribution for the next pitch. Because the mobility is constrained, the most likely pitches are predicted to be close to the most recent pitch. Moreover, because the most recent pitch is high in the tessitura, the most likely pitches are predicted to be somewhat lower.](image)

---

4. A simple test for homogeneity of variance is the rank correlation between the predictions of the regression equation and the absolute values of the residuals (Madansky, 1988). For the 176 sampled melodies, the median value of this rank correlation is minuscule ($r = -0.03$), indicating that the variances are close to homogeneous.
second-highest probability (18%) is assigned to the pitch A, which does in fact occur next in the melody.

Like prior definitions of pitch proximity, the new definition takes no account for scale structure. It assigns probabilities based entirely on the heights of pitches, regardless of where those pitches fit in a melody’s scale. In Figure 7, for example, the definition assigns the highest probability (19%) to the pitch A#, which does not fit the fragment’s D-major scale. If I were proposing a complete model of melodic structure, rather than an isolated definition for pitch proximity, I would probably combine the bell curve in Figure 7 with a second distribution representing the prevalence of different scale degrees. The key-profile distributions explored by Krumhansl (1990) and her colleagues might be suitable for this purpose.

Conclusion

To review, researchers have traditionally defined pitch proximity as a simple abundance of small pitch intervals. I advocate replacing this definition with a more refined alternative. Statistically, the proposed revision defines pitch proximity as a positive lag-one autocorrelation between successive pitch heights drawn from a quasi-normal distribution. More informally, the new definition is based on the premise that intervals are constrained both by a melody’s tessitura and by its mobility.

This new definition offers several advantages over the old one.

1. Whereas the old definition can predict only the sizes of melodic intervals, the new definition can predict their directions as well. In general, the new definition predicts that intervals will retreat from the extremes of the tessitura and approach the middle; intervals that start on high pitches will proceed downward, and intervals that start on low pitches will proceed upward.

2. Whereas the old definition predicts small intervals regardless of context, the new definition predicts that interval sizes will vary depending on the melody’s position in its tessitura. Near the middle of a melody’s tessitura, the new definition predicts small intervals, but near the extremes, the new definition predicts relatively large intervals.

3. The new definition subsumes two tendencies observed by music theorists: the tendency for a melodic peak to be approached by skip and the tendency for a skip to be followed by a change of direction. Although earlier work has presented these tendencies as independent rules of melodic structure, under the new definition they are seen as emerging from basic constraints on mobility and tessitura.
4. The new definition fits a statistical formalism that, given the first several notes in a melody, assigns a probability to each of the pitches that could occur next. More specifically, the new definition predicts that the probabilities for the next pitch will fit a normal distribution centered between the most recent pitch and the mean pitch of the melody. Both the peak and the width of this distribution are defined by statistical measures of tessitura and mobility.

In sum, compared with the old definition of pitch proximity, the predictions associated with the new definition are more precise and detailed, more sensitive to melodic context, and broader in their implications. In other words, the new definition enables a sharper and deeper description of melodic structure.

Although this definition of pitch proximity was developed to describe melodic structure, it might plausibly describe melodic psychology as well. A rule of pitch proximity is often used to account for such cognitive phenomena as melodic expectations, melodic preferences, and auditory streaming (Bregman, 1990; Carlsen, 1981; Dowling, 1967). In these cognitive applications, just as in descriptions of musical structure, pitch proximity has traditionally been defined simply in terms of interval size. In light of the present results, however, there is good reason to try an alternative definition. Perhaps melodic cognition, like melodic structure, could be described more precisely in terms of constraints on tessitura and mobility.5

References


---

5. I thank David Huron, David Temperley, action editor Adrian Houtsma, and two anonymous reviewers for thoughtful comments on drafts of this paper.
Redefining Pitch Proximity


