Difficulty Factors in Melodic Perception

Jeffrey L. Gillespie

As an instructor of twentieth-century aural skills for the past six semesters, I have become increasingly interested in the perception of melodies that do not fall clearly into the category of "tonal." I would like to guide students toward effective listening strategies, but I have difficulty determining which strategies are best for the students. Where and, more importantly, how should one begin with these tonally-comfortable students? At a special session of the 1987 Society for Music Theory conference, Elizabeth West Marvin posed similar questions:

> To what extent, then, does our perception of non-tonal materials rely upon 'default' techniques for tonal perception? Should aural skills curricula for non-tonal music build upon groundwork laid in tonal training (on 'altered' scales, scale-degrees, harmonies) or should they assume a completely separate atonal sound world?[1]

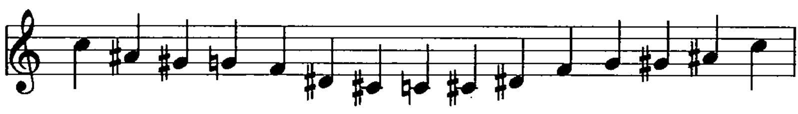
These questions are difficult to answer, and opinions vary considerably. Consider the published texts devoted to contemporary aural skills: Herder's *Tonal/Atonal*, which promotes a "tonally-referenced" approach, Adler's *Sight Singing*, with an interval-based format, and most recently, Friedmann's *Ear Training for Twentieth-Century Music*, featuring a set-class approach. Each text is designed for classroom use, and each is strikingly different. How can we choose? And *should* we choose one approach over another?

The few published studies that deal with "non-tonal" pitch perception have yet to provide many practical suggestions for the skills pedagogue. Several studies reflect biases on the part of the experimenter regarding what constitutes an atonal pattern. Example 1 presents two such patterns from earlier studies (Taylor, 1976; Dowling

[1]Complete citations for all sources mentioned are given at the end of the article.

EXAMPLE 1. Two melodies considered "atonal" in previous studies

Taylor (1976)



Dowling and Fugitani (1971)

and Fugitani, 1971). The Taylor melody may be heard as a C-Phrygian scale or an Ab-major scale with C as the top and bottom boundary, and the other pattern as centered around A or C. The studies often assign melodies to one of only two categories: tonal or atonal, with nothing in between. Since real twentieth-century music is not so cut-and-dried, these categories are of limited practical use. What one listener hears as atonal may indeed sound tonal to another.

Answers to these problems may be found from the students themselves, and the present pilot study represents a first attempt to do just that. The long-range objectives are to gain a better understanding of what makes a pattern difficult to hear and of how a listener most naturally perceives a pattern (intervallically, tonally, or in some other way). With this information, an instructor will be better equipped to consider more effective teaching strategies.

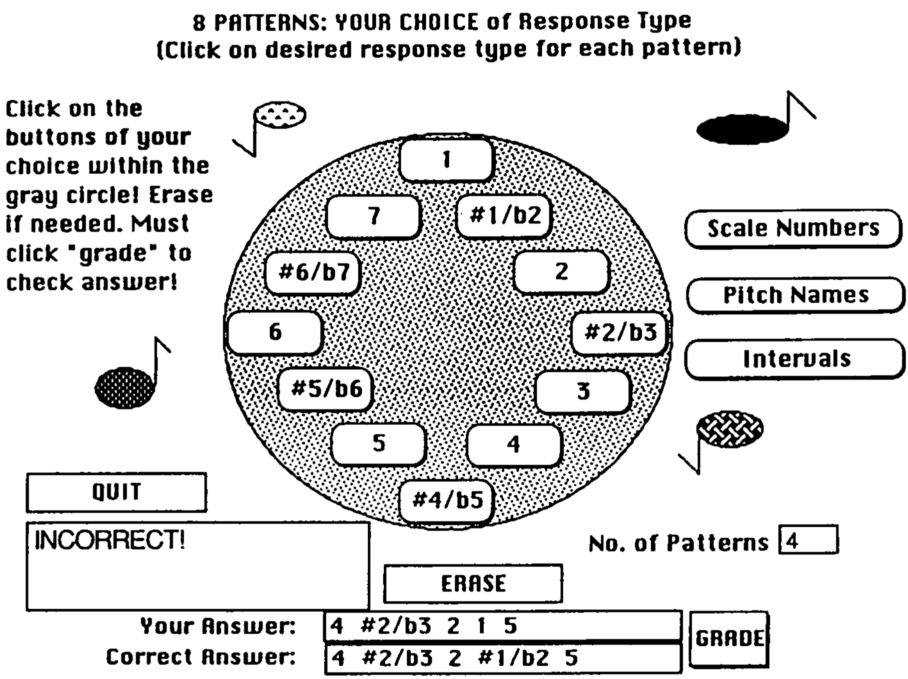
Several considerations served to guide the methodology. Since I was interested in obtaining a large body of data from subjects, and since I could make no prior predictions of what that data would reveal, I was drawn to ideas championed by Lincoln and Guba's *Naturalistic Inquiry*. According to these authors, a naturalistic study is "emergent" in design, with incoming data serving to guide subsequent stages of the project. Although I made certain decisions regarding the test format, I wanted subjects to respond based upon their "natural" tendencies when perceiving a pattern. To avoid biases in the construction and categorization of pitch patterns, I decided that all patterns for this study would be randomly generated from the chromatic scale.

Following the test, subjects were interviewed informally to glean additional feedback not revealed by the experimental data.

Six participants were chosen for their status as experienced listeners. Four subjects were Ph.D. theory students, three of whom had experience teaching aural skills. One subject was a theory faculty member, and the sixth was an undergraduate music major who had just completed the twentieth-century aural skills course with an A average. Additional information regarding these six participants is provided in Figure 6, which concludes this article.

The test itself consisted of 31 five-note patterns, generated randomly within a two-octave range around middle C, with no repeated pitches and a maximum range of a major seventh. A Macintosh computer served as the sound source. These 31 patterns were presented to the listener for transcription via a circular HyperCard grid, such as that presented in Figure 1. The continuous grid prevented the listener from having to displace octaves, as would have been necessary if the grid were horizontal.

FIGURE 1. Sample HyperCard screen taken from pilot study



Following the first hearing of a pattern, the subject was required to respond with a transcription. If the response was incorrect, a second hearing was given; the process was continued for a maximum of four tries. After the pattern was correctly transcribed or after four incorrect attempts, a new pattern was presented. Three possible response modes were available to the listener: scale degree numbers (or solfège, depending on the listener's preference), consecutive intervals, and pitch names. The test sequencing was designed to require some "forced" responses and some "choice": three practice patterns (one forced scale response, one forced interval response, and one forced pitch name response); eight with the listener's choice of response mode; four forced scale, four forced interval, and four forced pitch name; and eight more choice responses.

All incoming data was recorded as shown in Figure 2. In the top line of data, the first number (1) represents the pattern number. The following five numbers represent the actual pitches of the pattern presented, with middle C as 60. The next set of digits represents the listener's response in scale degree numbers, followed by the seconds taken for the response to be completed (18 seconds), concluding with "N" for an incorrect response.

FIGURE 2. Sample data listing, including exact pitches presented (middle C = 60), listener's response for each trial (indicating response mode as well), seconds taken to respond, and whether response was correct or incorrect (Y or N)

1,	64 68 70 63 67 ,	1 3 #4/b5 #1/b2 3 ,	18, N
1,	64 68 70 63 67 ,	1 3 #4/b5 #1/b2 4 ,	32, N
1,	64 68 70 63 67 ,	1 3 #4/b5 7 #2/b3 ,	46, Y
2,	60 61 67 62 69 ,	m2 P4 M3 P5 ,	16, N
2,	60 61 67 62 69 ,	m2 M3 P4 P5 ,	24, N
2,	60 61 67 62 69 ,	m2 P4 M3 P5 ,	25, N
2,	60 61 67 62 69 ,	m2 TT P4 P5 ,	29, Y
3,	65 66 62 64 71 ,	d#/eb E C F A ,	20, N
3,	65 66 62 64 71 ,	d#/eb E C D A ,	13, Y
4,	59 56 58 64 60 ,	2 7 #1/b2 5 3 ,	35, N
4,	59 56 58 64 60 ,	2 7 #1/b2 4 #2/b3 ,	21, N
4,	59 56 58 64 60 ,	2 7 #1/b2 5 #2/b3 ,	39, Y

FIGURE 3. Interval difficulty ratings, derived from confusion matrices of Blombach and Parrish (1988). The first figure represents "percentage correct" in their study, and figures in parentheses represent reciprocal values used in the present study to indicate interval difficulty level

	Ascending	Descending
P1	100 (0)	100 (0)
m2	85 (15)	82 (18)
M2	90 (10)	69 (31)
m3	78 (22)	60 (40)
M3	69 (31)	67 (33)
P4	86 (14)	69 (31)
TT	63 (37)	58 (42)
P5	80 (20)	75 (25)
m6	43 (57)	42 (58)
M6	56 (44)	57 (43)
m7	63 (37)	44 (56)
M7	62 (38)	36 (64)
P8	91 (9)	70 (30)

To determine what makes a pattern difficult to hear, rating systems can be designated to analyze the various features of the pattern. Items to consider are the number of direction changes, the range, the total consecutive interval content, and more complex judgments about interval difficulty level and the degree of "tonal fit." Interval difficulty ratings may be determined from previous interval studies, such as Blombach and Parrish (1988). Figure 3 presents figures from that particular study. The numbers without parentheses represent "percentages correct" taken from confusion matrices. For the current test, I used the reciprocal figures (in parentheses): the higher the number, the greater the expected difficulty in correctly identifying the interval.

Ratings for "tonal fit" may be determined from probe tone results of Krumhansl and Kessler (1982), presented in Figure 4A. Krumhansl's probe tone tests involve initially establishing a tonal context, in this case C major or C minor, through the playing of a scale fragment or a simple harmonic progression. Then a single pitch is presented from the chromatic scale; the listener is asked to rate how well that particular pitch "continues" or "fits within" the established tonal context. According to the figures in Figure 4A, C was selected as the pitch most closely related to the context, followed by the other pitches of the tonic triad.

While these probe tone ratings are presented in the key of C, they can easily be applied to any pattern placed within a scale system. Once a tonic is chosen for a five-note pattern, the Krumhansl values may be assigned to each of the pitches. A total of the five values provides quantified information on the tonal stability of the entire pattern within the chosen tonal placement. Since the original figures increase with tonal strength, I again chose to use reciprocal values, so that higher numbers indicate an increase in non-tonal strength. The higher the total of all scale degree values in the pattern, the more difficult that pattern may be to perceive according to a particular tonal placement. For any given pattern, there are twelve different tonic placements from which a listener may choose, including tonics that are not present in the pattern itself. Since a listener is expected to focus on one particular tonal center if responding by the scale system, it is the lowest figure (or "most tonal" placement) that is most significant. For illustration, Figure 4B presents a sample pattern, with calculated values for "non-tonalness" for each of the twelve possible tonic placements, based upon the "C Major Context" figures of Krumhansl and Kessler.

Example 2 presents a sampling of patterns selected from the original test. The patterns include a variety of difficulty factors for comparison. For instance, for the third pattern listed (C#-F-G-F#-A), the first rating of 11 represents the total range of the pattern—eleven semitones from G up to F#. The second number indicates three changes of direction within the pattern, following the initial ascent from C# to F. The total interval content, or the semitone sum of all the consecutive intervals, is 34. The interval difficulty rating, based on the figures of Blombach and Parrish (Fig. 3), is 168. And finally, 15.27 represents the "most tonal" placement, according to the Krumhansl figures from Figure 4: if C# is chosen as tonic, then the sum of the corresponding values of the scale degrees in the pattern is 15.27. This high value indicates that the pattern is quite non-tonal in comparison with other patterns presented in the test.

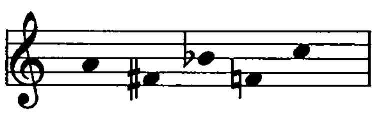
With definitive values for these difficulty factors, correlation studies can be run against subjects' response success to determine which factors most influence perceptual accuracy. Though a larger group of subjects is needed to attempt such correlations, several observations may be made with the six participants of the current study (Fig. 5).

FIGURE 4.

4A) Probe tone ratings from Krumhansl and Kessler (1982), taken from Table 2.1 of Krumhansl (1990). Since these figures increase with greater tonal stability, the present study used reciprocal values (in parentheses) to correlate increase in value with increase in tonal instability

Tone	C Major Context		C Minor Context	
C	6.35	(.65)	6.33	(.67)
C#/Db	2.23	(4.77)	2.68	(4.32)
D	3.48	(3.52)	3.52	(3.48)
D#/Eb	2.33	(4.67)	5.38	(1.62)
E	4.38	(2.62)	2.60	(4.40)
F	4.09	(2.91)	3.53	(3.47)
F#/Gb	2.52	(4.48)	2.54	(4.46)
G	5.19	(1.81)	4.75	(2.25)
G#/Ab	2.39	(4.61)	3.98	(3.02)
A	3.66	(3.34)	2.69	(4.31)
A#/Bb	2.29	(4.71)	3.34	(3.66)
B	2.88	(4.12)	3.17	(3.83)

4B) Sample pattern with non-tonal strength values calculated for all twelve possible tonic placements, based upon the original figures of Krumhansl and Kessler (1982). According to the values, the most non-tonal placement would be with E as tonic, and the most tonal placement would be with F as tonic.



Tonic

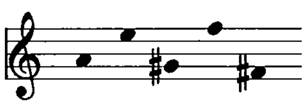
C = 16.09 (3.34 + 4.48 + 4.71 + 2.91 + .65 = 16.09)

C#	= 17.6		
D	= 18.42	G	= 19.93
D#	= 17.82	G#	= 18.96
E	= 20.29	A	= 18.04
F	= 12.76	A#	= 14.71
F#	= 16.54	B	= 19.89

Item B indicates the percentage of errors for each of the four consecutive intervals of all response patterns; participants were definitely most accurate at the beginnings of responses and least accurate at the ends. Additional observations may be made for those patterns labeled by scale degree numbers. Figure 5C indicates that in 43.66% of the scale responses the first note was chosen as tonic. This percentage is interesting when compared with the percentage of first pitches that represent Krumhansl's "most tonal" tonics: 19.72%. When placing patterns within a tonal system, listeners chose the strongest "Krumhansl tonal fit" in only 38.03% of the responses. This disappointing percentage does little to support Krumhansl's original figures (see Fig. 4). If, however, the "top four" strongest tonal placements are considered for each pattern, based on Krumhansl's values, the percentage is much higher: in 83.1% of the responses, listeners chose one of these four tonics.

EXAMPLE 2. Sample patterns taken from pilot test, with "difficulty factor" ratings (Range / Direction Changes / Interval Total / Interval Difficulty / Non-Tonalness) and generalized categories for comparison (H - High, M - Medium, L - Low)

Highest Overall Ratings | Ran/Dir/ITot/IDif/Non-Tonal



11 / 3 / 35 / 186 / 14.47 (1=E)
(H H H H M)

10 / 3 / 26 / 165 / 15.88 (1=B)
(H H H H H)

11 / 3 / 34 / 168 / 15.27 (1=C#)
(H H H H H)

Lowest Overall Ratings

7 / 1 / 12 / 77 / 13.66 (1=Eb)
(L L L L M)

8 / 2 / 14 / 100 / 12.9 (1=Eb)
(M M L L L)

EXAMPLE 2. (cont.)

"Featured" Factors — Ran/Dir/ITot/IDif/Non-Tonal


11 / 1 / 12 / 61 / 12.47 (1=A)
(H L L L L)

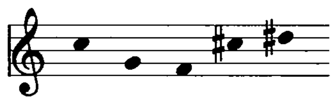

4 / 1 / 6 / 77 / 16.23 (1=F)
(L L L L H)


7 / 3 / 14 / 84 / 13.27 (1=F)
(L H L L L)


9 / 1 / 17 / 161 / 13.09 (1=A)
(M L M H L)


10 / 3 / 28 / 157 / 12.9 (1=A)
(H H H H L)

More "Non-Tonal" Values


10 / 1 / 17 / 129 / 14.8 (1=Ab)
(H L M M M)

(Most "tonal" tonics = Ab [14.8], C [14.81], Eb [14.84],
F [15.3]; least tonal tonic = E [21.51])


7 / 3 / 19 / 122 / 12.76 (1=F)
(L H M M L)

(Most tonal tonics = F [12.76], Bb [14.71], C [16.09],
F# [16.54]; least tonal tonic = E [20.29])

Finally, a few personal observations about the participants. All but one listener averaged between 26 and 33 seconds for each response, with one averaging 58 seconds. Most interesting is the response preferences when participants were given a choice. One respondent answered by interval every time. Another consistently chose scale degree numbers and verbally expressed annoyance at being forced to re-

FIGURE 5. Statistical information

A)

Participant	# Correct 1st Try	Ave. Incorrect Tries per Pattern	Ave. Time per Pattern
#1	0	2.25	33.06"
#2	10	1.25	28.98"
#3	5	1.86	33.00"
#4	11	1.46	28.08"
#5	6	1.57	26.33"
#6	3	2.25	57.70"

B) % OF ERRORS (by interval placement):

1	2	3	4
10.6%	22.5%	32.8%	34.1%

C) % OF RESPONSES WITH 1ST NOTE CHOSEN AS TONIC: 43.66%
(% OF PATTERNS WITH 1ST NOTE AS "MOST TONAL" TONIC, ACCORDING TO KRUMHANSL: 19.72%)

D) % OF RESPONSES WITH "MOST TONAL" TONIC CHOSEN: 38.03%
% OF RESPONSES WITH ONE OF THE "TOP FOUR" TONICS CHOSEN: 83.10%

spond by interval during certain portions of the test. The other four listeners exhibited a distinct preference for responding by scale degree, with occasional interval response. Since background and training are likely to play a part in such response choices, brief profiles of each participant are included in Figure 6.

For the next stage in this study, I plan to expand data collection by testing a greater number of listeners, at least 30, using the same method of randomly generated patterns. With a larger amount of data, linear regression analysis may be used to evaluate each individual difficulty factor regarding its degree of influence on listeners' perceptual accuracy. Subsequently, I hope to test listeners with actual melodies from the literature. With information gathered from studies such as this, aural skills instructors will be better equipped to organize melodic examples for dictation purposes and to achieve a clear progression from one difficulty level to the next during the course of a semester. With more subjects tested, a better understanding would be gained regarding the natural tendencies of listeners: Is there a general preference for the scale system approach no matter what the tonal fit of patterns may be? Is each listener different in preference, perhaps based on prior experience? Which listeners achieve the highest accuracy levels, those who favor a scale system approach, an interval approach, or a flexible approach to listening? With more information on listeners' natural tendencies and corresponding accuracy levels, we as teachers can better guide students to reach their potential as listeners, in ways that are most comfortable to the students.

FIGURE 6. Profiles of the six participants

Participant #1
Major Instrument: Voice Level: Senior
Aural Skills Training: Scale degree numbers; completed undergraduate sequence (including non-tonal aural)
Teaching: None
Music for Pleasure Listening: Primarily Baroque
Test response preference: Intervals (consistently)

Participant #2
Major Instrument: Piano Level: Ph.D. theory student
Aural Skills Training: Solfège; private aural skills training, no non-tonal aural training
Teaching: Sophomore level aural skills, twentieth-century aural skills, graduate aural review
Music for Pleasure Listening: Contemporary "classical"
Test response preference: Solfège, occasional interval

Participant #3
Major Instrument: Viola Level: Theory faculty
Aural Skills Training: not available
Teaching: Twentieth-century analysis ca. fifteen times, twentieth-century aural skills three-four times
Music for Pleasure Listening: Jazz, chamber music
Test response preference: Scale numbers, occasional interval

Participant #4
Major Instrument: Saxophone Level: Ph.D. theory student
Aural Skills Training: Solfège and scale numbers; no non-tonal aural skills
Teaching: Freshman and sophomore ear training for four+ years
Music for Pleasure Listening: Gospel
Test response preference: Scale numbers (consistently)

(continued)

FIGURE 6. (cont.)

Participant #5
Major Instrument: Piano/Organ Level: Ph.D. theory student
Aural Skills Training: Scale numbers; no non-tonal aural skills
Teaching: Twentieth-century analysis twice, twentieth-century aural skills six times
Music for Pleasure Listening: Classical, jazz, pop/rock, country, gospel, show tunes
Test response preference: Scale numbers, occasional interval

Participant #6
Major Instrument: Piano Level: Ph.D. theory student
Aural Skills Training: Two-three semesters of basics, no non-tonal
Teaching: Twentieth-century analysis, freshman and sophomore written theory
Music for Pleasure Listening: Twentieth-century Russian and English, Bartók, classical, rock, Celtic-folk
Test response preference: Pitch name, some scale numbers

REFERENCES

Adler, Samuel. *Sight Singing.* New York: W.W. Norton, 1979.

Blombach, Ann K., and Parrish, Regena T. "Acquiring aural interval identification skills: Random vs. ordered grouping." *Journal of Music Theory Pedagogy* 2/1 (1988): 113-131.

Butler, David. "Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry." *Music Perception* 6/3 (1989): 219-241.

Dowling, W.J., and Fugitani, D.S. "Contour, interval, and pitch recognition in memory for melodies." *Journal of the Acoustical Society of America* 49 (1971): 524-531.

Friedmann, Michael L. *Ear Training for Twentieth-Century Music.* New Haven: Yale University Press, 1990.

Herder, Ronald. *Tonal/Atonal.* Rev. ed. New York: Continuo Music Press, 1977.

Krumhansl, Carol L., and Kessler, E.J. "Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys." *Psychological Review* 89 (1982): 334-368.

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

Lake, William E. "The Perception of Melody: Implications of Recent Research." M.M. thesis, Indiana University, 1982.

Lincoln, Yvonna S., and Guba, Egon G. *Naturalistic Inquiry.* Beverly Hills, CA: Sage Publications, 1985.

Marvin, Elizabeth West. "Music perception and theories of non-tonal music." Paper presented at the meeting of the Society for Music Theory, Rochester, New York, November 1987.

Taylor, Jack A. "Perception of tonality in short melodies." *Journal of Research in Music Education* 24 (1976): 197-208.