Research on Tonal Perception and Memory: What Implications for Music Theory Pedagogy?[1]

Elizabeth West Marvin

Introduction

Let us begin with a few straight-forward questions: why should music theorists read psychology journals? What can psychologists who are nonmusicians or amateur musicians tell professionals about music hearing and learning? And what can the results gleaned from highly controlled laboratory experiments have to do with our "real-world" theory classrooms? For teachers of music theory who have pondered these questions, but have not yet found time to sit down with the appropriate books and journals, be assured that music theorists are addressing these questions (and more) in print. If you know where to look, you will find music theorists thoughtfully critiquing the cognitive science literature, interpreting empirical research for the classroom, and implementing their own experiments based on issues that have arisen in their teaching.

To those who are new to this research area, however, the sheer number of experimental studies on music perception and memory may seem daunting—hence the reason for this article, which is intended to serve as an introduction to the field for classroom teachers of music theory. Because the discipline's purview is so broad and diverse, this discussion centers around a recommended read-

[1]An early version of this paper was presented as the Keynote Address for the Texas Society for Music Theory's 1994 annual meeting at Lamar University (Beaumont, TX).

ing list, intentionally limited just to the areas of tonal perception and memory—critical elements for undergraduate theory teaching. The articles cited examine patterns of interval recognition and confusion, cognitive strategies for melodic recall, sight singing, and dictation, and perception of tonal hierarchy, modulation, and large-scale formal design. Topics such as rhythmic perception, aesthetic response, expressive timing in performance, and perception of nontonal music have likewise generated a considerable body of empirical research, but a discussion of their results and pedagogical implications is beyond the scope of a single essay.

The selected bibliography at the end of this article is divided into four parts; this organization models both the ordering of our discussion and the order in which interested readers might tackle the list. Although the bibliography is organized topically, books and essays are numbered consecutively throughout to facilitate their citation in the body of this article; Arabic numerals enclosed in parentheses will be used to refer to specific entries. The bibliography's first part recommends three good overviews of the field, from which the reader will quickly discern the scope of the discipline.[2] Part two cites a number of experiments concerned with the mental representation of melodies. This is an area of research with clear implications for the classroom, and several articles by music theorists attest to this fact. Part three focuses upon the perception of tonal hierarchy and closure, an area of research in which music theorists have taken some psychologists to task. These studies ask some very basic questions with important implications for theory instruction: for example, can we really hear hierarchical structure in music? Does the *Ursatz* have perceptual validity? Is it appropriate to teach analytical methods that reveal relationships many listeners may not hear? Part four is organized around the issue of individual differences: what does music-cognitive research tell us about the differences among our music students? What, for example, should teach-

[2]Music libraries will not carry all of the journals represented in the bibliography; some are specialized journals in psychology. Interested readers will need to use a university library or rely upon inter-library loan. However, the three excellent texts cited in the first part of the bibliography ought to be acquired by any good music library, as should the interdisciplinary journal *Music Perception*.

ers do with absolute pitch (AP) students— simply exempt them from sight singing classes by exam? Are there different music-cognitive strategies between AP and non-AP students, between composers and performers, even between male and female musicians?

Interaction between music-theoretical and music-cognitive research is still in its early stages, but it is growing. This growth may be seen in the increasing number of music theorists presenting papers at the annual meetings of the Society for Music Perception and Cognition (over a dozen at the 1993 annual meeting in Philadelphia),[3] and by events such as the Special Session at the 1992 Society for Music Theory meeting, where prominent psychologists Edward Carterette, Jay Dowling, and Carol Krumhansl shared the podium with music theorists Patrick McCreless, Joseph Dubiel, and Janet Hander-Powers.[4] In many ways, the interaction between disciplines still represents an uneasy alliance, however; thus some articles by music theorists in the bibliography cite the results of music-cognitive experimentation to support specific pedagogical recommendations, while others criticize specific experimental studies for flaws in their methodology or in the interpretation of their data.

Research on Melodic Perception and Recall

We begin our discussion on the positive side of the uneasy alliance, with articles (#14-16 in the bibliography) by Gary Karpinski, Bill Lake, and Steve Larson—all writing in the *Journal of Music Theory*

---

[3]Presenters included Robert Gjerdingen, Edwin Hantz, David Huron, John Kelleher, Steve Larson, Fred Lerdahl, Justin London, Elizabeth West Marvin, Richard Parncutt, Jay Rahn, Burton Rosner, and William Thomson.

[4]At least one of these papers has now appeared in print: Carol Krumhansl's "Music Psychology and Music Theory: Problems and Prospects" in *Music Theory Spectrum* 17/1 (1995): 53-80. Of particular interest are the article's first eight pages, which present a "schematic history of the psychology of music from the point of view of how it has been influenced by music theory" (p. 80). The remainder of the article discusses ways in which the basic tenets of Eugene Narmour's implication-realization model for melodic expectancy can be fashioned into testable hypotheses, and Krumhansl reports the results of an experiment she designed and carried out to test these hypotheses.

Pedagogy—which draw upon research in music cognition to support the pedagogical approaches they use in their classrooms. To some degree, these articles (and the present one) answer the challenge posed by David Butler and Mark Lochstampfor to "synthesize experimental results from the research area of music cognition so that they may be applied directly to aural training."[5] Karpinski, Lake, and Larson all share a pedagogical interest in how students represent melodies in memory. Teachers generally must infer students' mental representations by observing their successes or failures in melodic dictation and sight singing. In the United States, most music theorists impose a scale-degree based strategy for melodic representation upon their students when teaching these skills, as evidenced by the prevalence of movable-do solfège and scale-degree sight singing instruction over fixed-do solfège, which instead teaches pitch-class identification.[6] Steve Larson's article (#16) strongly advocates movable-do solfège, with do-based minor, as the best pedagogical model of scale-degree function, and he supports his argument by drawing upon psychological experimentation and principles of Gestalt psychology.

It may come as some surprise to music theorists that the mental representation of melodic pitches in relation to a key center has been an issue of considerable controversy and experimentation in music-cognitive research over the years. At the root of these experiments is the very basic question cited in Figure 1: just what is it that listeners remember when they hear a melody? Is it a pitch-class succession, a contour plus scale type, an interval succession, scale degrees in a key, or some combination of these? The first of these possibilities is that listeners remember melodies as collections of

---

[5]David Butler and Mark Lochstampfor, "Bridges Unbuilt: Aural Training and Cognitive Science," *Indiana Theory Review* 14 (1993): 6.

[6]For documentation, see Randall G. Pembrook and H. Lee Riggins, "'Send Help!': Aural Skills Instruction in U.S. Colleges and Universities," *Journal of Music Theory Pedagogy* 4/2 (1990): 231-241. The authors note (p. 236) that of 518 responses, 57% of respondents use some type of movable system (either numbers or movable-do solfège), while only 19% use a fixed system (either fixed-do or letter-name singing, with or without inflections).

pitch classes. Pitch-class recognition would explain the ease with which listeners recognize untransposed melodic repetitions or melodies repeated an octave higher or lower. It would not account, however, for the ease with which listeners recognize exact transpositions, or the difficulty listeners experience when asked to recognize familiar melodies in which random pitches have been displaced by an octave, thus retaining their pitch-class identity but possibly violating the contour of the melody. This octave displacement effect has been clearly demonstrated by Diana Deutsch in her experiments on pitch and pitch class.[7] These experiments are briefly summarized by Rita Aiello in her introduction to W. Jay Dowling's "Melodic Contour in Hearing and Remembering Melodies" (#9).[8]

The pitch-class model of melodic representation cannot explain the ease with which listeners recognize and sing tonal melodies transposed to other keys, especially distant keys that share few common pitch classes. This question of melodic recognition after transposition was investigated by James Bartlett and Jay Dowling in 1980

Figure 1. How do we represent melodies in memory?

Strategy	*"Are You Sleeping"*
1. Pitch-class succession	C-D-E-C
2. Contour plus scale type	<++-> plus major scale
3. Interval succession	+M2 +M2 -M3
4. Scale degrees in a key	Major key: 1-2-3-1 *or* do-re-mi-do
5. Some combination of these	

---

[7]In Deutsch's early publications, the terms "pitch height" and "chroma" are used for "pitch" and "pitch class," respectively. Deutsch summarizes her findings in "The Processing of Pitch Combinations," in *The Psychology of Music*, ed. Diana Deutsch (New York: Academic Press, 1982).

[8]An aural example based on Deutsch's octave-displaced melodies can be heard on the compact disc that accompanies David Butler's fine book, *The Musician's Guide to Perception and Cognition* (New York: Schirmer Books, 1992). This demonstration could be used effectively in theory classes to illustrate the aural difference between pitches and pitch classes in a musical context. Butler's text comes not only with a compact disc of

(#5); the experimenters asked listeners to distinguish between exact transpositions of unfamiliar melodies (called "targets") and intervallically altered replications of these melodies (called "lures"). Listeners more often confused lures with targets when heard in distant keys than when they were heard in closely related keys. The authors concluded that "the results support the psychological reality of key distance and are consistent with both musical and non-musical-auditory theories of its effects."[9] Annie Takeuchi and Stewart Hulse (#12) subsequently refuted this finding in three different experimental contexts; they pointed out that if key-distance effects were a psychological reality, then this would imply that most listeners have some degree of absolute pitch, and that they "encode" pitch-class note names in memory. Aural skills teachers would suspect that this is not true; this intuition is confirmed by Takeuchi and Hulse, as well as by Croonen and Kop (#6).

In some of Dowling's early work, he posited the melodic representational scheme given as the second option of Figure 1: that melodic contour is an abstraction that can be retained independently of pitches or intervals. In his two-component model for memory of melodies, tonality was posited as a kind of framework upon which contour may be "hung"; thus listeners would actively reconstruct melodies in memory from scale information plus contour. This model would explain Dowling's finding that listeners have difficulty distinguishing between a tonal melody and a same-contour lure that begins on a different scale degree of the same key, since the two melodies represent the same scale and the same contour, just placed in a different position along that scale.[10] More recently, however, Dowling has abandoned this idea.[11] This change of opinion

examples, but also with an instructor's manual (available from the author on floppy disc by request).

[9]James C. Bartlett and W. Jay Dowling, "Recognition of Transposed Melodies: A Key-Distance Effect in Developmental Perspective," *Journal of Experimental Psychology: Human Perception and Performance* 6 (1980): 501.

[10]W. Jay Dowling, "Scale and Contour: Two Components of a Theory of Memory for Melodies," *Psychological Review* 85 (1978): 341-354.

[11]W. Jay Dowling, "Melodic Contour in Hearing and Remembering

was reached in part because of more recent research showing that tonal context affects memory for contour. For example, in a 1991 experiment, Dowling asked listeners to discriminate between exact transpositions of a target melody and same-contour or different-contour lures. He used melodies of three types (strongly tonal, weakly tonal, and atonal), and he systematically varied the delay time between presentation of the stimulus and the subjects' response. Figure 2 summarizes his results in graphic representation.[12] After short delays, subjects discriminated targets from different-contour lures better than they did from same-contour lures in all three conditions: strongly tonal, weakly tonal, and atonal. The change in contour seemed to be the primary cue by which subjects discriminated between stimuli. By increasing the delay time, however, this difference tended to disappear for the strongly tonal melodies. With time, performance in the tonal condition improved for discrimination of same-contour lures. For atonal melodies, however, the longer delay time resulted in poorer discrimination. Dowling concluded that "listeners' ability to reject a different contour test item depended on tonal context, and not just on its difference of contour."[13] This result implies that contour and scale type are not independent components in memory for melody. It means too that the longer delay time actually made the task easier in the tonal condition; it gave subjects time to encode and process information other than simple

---

Melodies," in *Musical Perceptions*, ed. Rita Aiello with John Sloboda (New York: Oxford University Press, 1994), 186. Though Dowling has abandoned this idea, others still espouse it. Witness, for example, the statement by Mariko Mikumo ("Encoding Strategies for Tonal and Atonal Melodies," *Music Perception* 10 (1992): 28): "These interpretations suggest that contour or shape and pitch or interval are independently handled as different features and, moreover, suggest that contour or shape recognition is dominant over pitch or interval recognition." Perhaps contour plays a stronger role in atonal melodies, thus accounting in part for her results.

[12]Figure 2 is reproduced from Dowling "Melodic Contour," 187. It summarizes results he reported in "Tonal Strength and Melody Recognition after Long and Short Delays." Perception and Psychophysics 50 (1991): 305-313.

[13]Ibid., 186.

contour. It is Dowling's hypothesis that contour information is the primary strategy for discrimination after short delays, and that interval information becomes available after longer delays; the ultimate result of this change in strategy is that intervals, in his opinion, are of primary importance in long-term memory for melodies.[14]

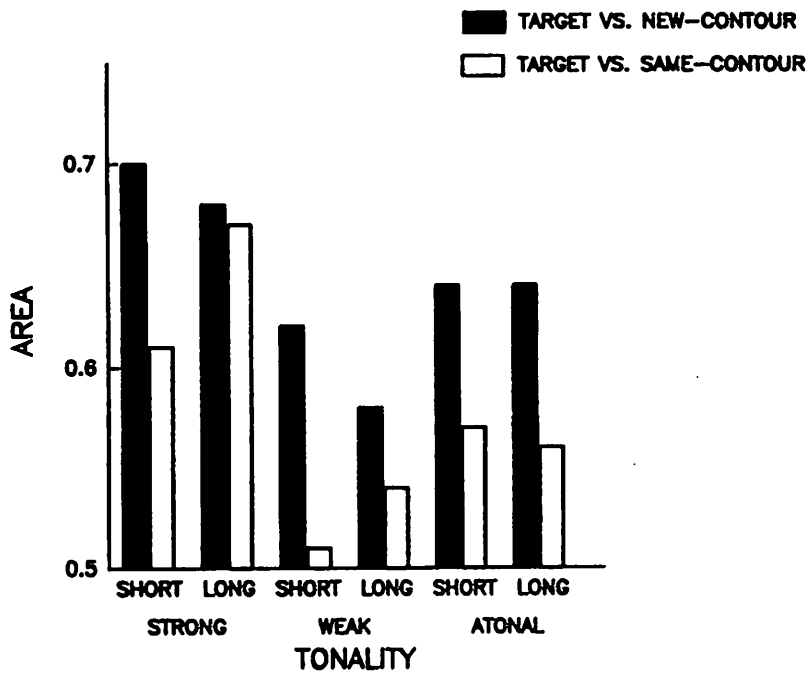
In evaluating the results of the experiment shown in Figure 2, Dowling does not consider that the two different classes of melodies (tonal and atonal) might necessitate two different types of listening strategies. Music theory teachers will suspect that subjects' differential ability to discriminate between target and lure in the two conditions probably indicates a difference in cognitive strategy that is dependent upon the strength of each melody type's tonal context. They might hypothesize further that subjects' ability to discriminate better in the tonal condition suggests that a different and more effective strategy has been used: namely, scale-degree information. To test this hypothesis, I designed and ran an experiment (#34) using three subject groups—absolute pitch musicians, non-AP musicians, and nonmusicians. The experiment's design was similar to Dowling's (discrimination between transposed targets and intervallically altered lures, tonal and nontonal) except that contour remained the same from target to lure and therefore could not be used at all as a discrimination cue. For all three subject groups (even nonmusicians), for both types of melody (even nontonal), subjects were able to discriminate at above chance levels, demonstrating that in the absence of contour cues, subjects were able to make use of other musical information in discriminating between target and lure. There were, however, differences in all three subject groups between listeners' performance in the tonal and nontonal conditions, with performance in the tonal condition significantly better.

---

[14]Ibid., 183. Curiously, Dowling seems to contradict this position just two pages earlier, when he says (p. 181): "Since interval patterns most obviously remain constant through transposition, a good case can be made that they are the most fundamental. Nevertheless, I believe that melodies are remembered as ordered sets of pitches. The pitch framework provided by the scale temporarily fixes a set of pitches as landmarks in terms of which the pitches of the melody are organized and understood." I suspect that Dowling misspeaks here, and that by "pitch framework provided by the scale" he means scale degrees, not pitches.

Figure 2. Results from Dowling (1991)

Short and Long Response Time
Melodies Strongly Tonal —> Atonal:



This finding suggests either that listeners are better at the tonal task simply because they have more experience listening to tonal music, or that listeners are better at this task because they use a different and more effective method of encoding the melodic information in memory. Both alternatives may contribute to this finding. Even though both scale-degree and interval representations remain the same when melodies are transposed, scale-degree patterns are easier to remember because of "chunking" (that is, for example, grouping triadic embellishments into one "item" in memory, rather than several discrete units). Listeners using this strategy need not remember a string of unrelated intervals; rather, they remember the leaps as patterns within the underlying context of a familiar harmonic progression of just two or three chords. Stepwise patterns may also be remembered in relation to triadic arpeggiation, as stepwise filling in between chordal members, for example. The fact that this type of chunking is not available in nontonal contexts accounts, in part, for musicians' poorer performance in the nontonal condition, since listeners in that condition had to remember a succession of tonally unrelated intervals.

Musicians are conditioned to chunk melodies by harmonic function even when they practice arpeggios on their instruments—a common technical exercise. Although this cognitive strategy may not be explicitly articulated to music students in their early years of instrumental study, continued involvement with the harmonic language of tonal music in performance, sight reading, and memorization enables experienced performers not only to infer a tonal center from a melody as it is heard, but also to group leaps within triads and even to infer implied harmonic progressions.[15] This ability to chunk melodies by harmonic function would account, in part, for the large differences found between musicians' and nonmusicians' discrimination in the tonal condition of the experiment just discussed—a difference which narrowed considerably in the nontonal condition, where chunking by function was not an option.[16]

---

[15]Research has also shown that musicians who effectively "chunk" by harmonic function are better sight readers. See, for example, John A. Sloboda, "Experimental Studies of Music Reading: A Review" in *Music Perception* 2 (1984): 222-236. Among the conclusions these researchers have reached are that good sight readers have better visual memories for notation, that they sense structural configurations in music, and that what they read is analyzed for its musical significance before the performer forms the appropriate motor commands.

[16]Investigators routinely divide subjects into "musician" and "nonmusician" categories, using the number of years of applied music study as a primary determinant. They frequently find significant differences in performance between these two subject groups, and some hypothesize that entirely different cognitive processes are at work in the two groups. Yet, consider that some of our music students (typically singers) do not begin applied music study until adolescence. If the number of years of instrumental study is used as the distinguishing criterion for subject grouping, these late-starting music students might fall

In related research, Lola Cuddy (#7) asked subjects to discriminate between a transposed target melody and a lure in which just one pitch had been altered by a semitone from a true transposition; in a later replication of the experiment, she altered the lure's one pitch by only a quarter tone! Cuddy's melodies were systematically constructed to conform to five different tonal strengths, ranging from strongly tonal to atonal, and she controlled for contour type. In all three subject groups, Cuddy found that discrimination of target melodies from lures was highly correlated to tonal strength. This was true even in the experiment using the quarter-tone alteration. Cuddy attributed the high level of discrimination in the strongly tonal condition to the fact that "perception of melody involves a sense of progression among notes . . . [which] is itself a sense of regular ebb and flow from a central nodal point (the perceived tonic)."[17] Thus, without using our terminology, Cuddy describes not a pitch-class nor an interval strategy, but rather a scale-degree strategy.

Cuddy's results show that the "ability to detect and implement harmonic structure exists in the average non-musician and this is a finding . . . that deserves further exploitation in adult music education."[18] Thus, the ability to make use of implied harmonic structure is available to most listeners; it is our job as teachers to enhance and reinforce it. One way to strengthen this ability is to employ pedagogical methods that allow students to perceive hierarchical relationships among scale degrees, and to hear scale-degree patterns that imply harmonic function, rather than methods that emphasize pitch-class identification or interval-based strategies for music hearing, singing, and dictation. As Cuddy's research showed, even listeners' ability to detect errors in intonation is affected by their ability to perceive underlying harmonic function within melodies.

---

into the "nonmusician" group, and might tend toward different cognitive strategies than their peers who began instrumental study at a younger age.

[17]Lola Cuddy, "On Hearing Pattern in Melody," *Psychology of Music* 10 (1982): 9.

[18]Ibid., 3.

Let us return for a moment to the interval hypothesis for representing melodies in memory. This approach is taken by one prominent sight-singing text: Samuel Adler's *Sight Singing: Pitch, Interval, Rhythm.* Adler contends that "if you sing by intervals at all times, you will be able to read music that is tonal, modal, pantonal, nontonal, or a combination of all these means of organization."[19] In an experiment designed to tease apart scale-step strategies from interval strategies among different subject groups, Dowling (#8) presented listeners with stimuli like those shown in Figure 3: a chordal context followed by a novel melody whose scale-step identities were to be interpreted in the context of the key previously presented.[20] As this figure shows, the chordal context was either a closed progression (I-IV-V-I) in the key of the melody, or an incomplete progression (I-IV-cadential 6/4-V7) in the key of the subdominant, thus suggesting reinterpretation of the given melody's scale degrees in a new key. The melody was also followed by a tonic chord confirming the key. Listeners were asked to distinguish between exact transpositions of the target melodies and lures in which one note had been changed. In surprising results, Dowling found that inexperienced listeners discriminated between targets and lures in the shifting tonal context better than moderately experienced musical subjects.

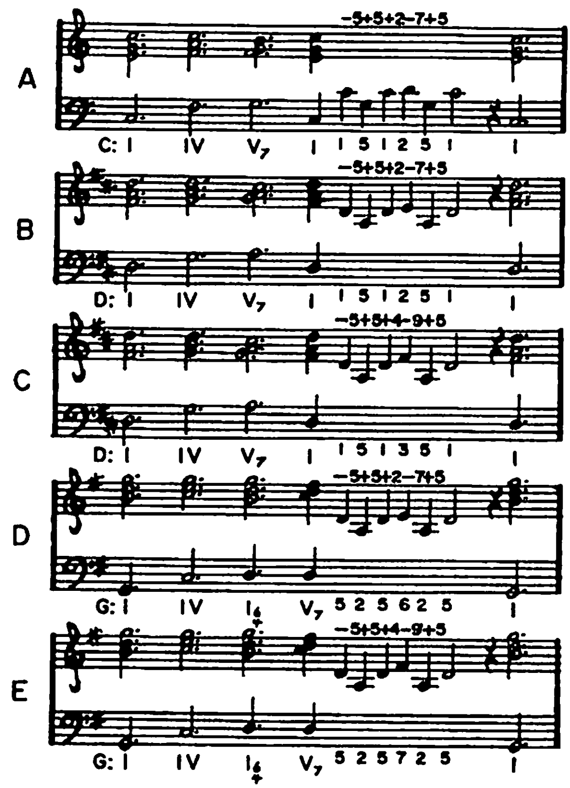
Figure 4 shows Dowling's conclusions. He attributed his finding to the use of an interval-based listening strategy by inexperienced listeners, which would be unaffected by the change in context. Listeners with moderate musical experience, he hypothesized, were more likely to remember the melodies in their scale-degree context; shifting tonal context would therefore cause confusion errors among listeners in the moderately experienced subject group. Professional musicians, while "capable of scale-step representation, were able to use a flexible memory-retrieval system to avoid errors

---

[19]Samuel Adler, *Sight Singing: Pitch, Interval, Rhythm* (New York: W. W. Norton & Company, 1979), x.

[20]Figure 3 is reproduced from "Context Effects on Melody Recognition: Scale-Step versus Interval Representations," *Music Perception* 3 (1986): 284.

Figure 3. Changing tonal contexts from Dowling (1986)



Examples of stimuli in Experiment 1: (A) A novel melody introduced with chordal context ending with the tonic (I) chord; (B) a same-context transposition of A; (C) a same-context imitation of A; (D) a different-context transposition of A; (E) a different-context imitation of A. The Roman numerals under the staves indicate the chord labels in the context; the small numbers above the melodies indicate the interval pattern in semitones; and the small numbers below the melodies indicate the scale-step values of the pitches. Note that transpositions preserve interval pattern exactly, and that the chordal context determines the initial scale-step values.

with changed context."[21] This, of course, is the point to be made in an effective theory pedagogy. Students need to develop a flexible memory retrieval system that is capable of both interval and scale-degree information. Absolute pitch musicians, for reasons to be discussed later, also need to develop a flexible system—one that uses not only the pitch-class information readily available to them, but also interval and scale-degree information that they often must be trained to use.

William Lake's article on "Interval and Scale-Degree Strategies in Melodic Perception" (#15) cites the Dowling study just described as his point of departure; Lake notes a tendency among his theory students to espouse one strategy strongly over the other, in spite of the fact that each system has some built-in problems:

> Neither approach is ideal for all situations. A scale-degree strategy is most suitable for tonal melodies. A typical problem for an interval thinker is to misjudge a leap and then remain off by a step; thinking scale degrees, however, allows quick reorientation after such a mistake. On the other hand, an interval strategy works best for atonal melodies and passages in tonal melodies that are modulatory or tonally vague. A typical mistake for a scale-degree thinker, who is not careful about specific intervals, is to alter notes to fit into some key (either imagined or no longer in force).[22]

Lake's classroom experience does not, however, fit with Dowling's "developmental" description of the interval strategy for less-experienced musicians and the scale-degree strategy for moderately experienced. In Lake's homogeneous group of students, both strategies exist. Indeed, he finds Dowling's developmental succession unconvincing; he notes "it seems unlikely that students would forget how to listen intervallically while they are learning scale-degree representations, only to recover that ability as professionals."[23] Lake suggests that teachers test students to discover whether they tend to hear by interval or by scale degree, then gear specific exer-

---

[21]Ibid., 281.

[22]Lake, 57-58.

[23]Lake, 58.

Figure 4. Melodic recognition strategies from Dowling (1986)

Inexperienced listeners	——>	Interval series
Moderately experienced	——>	Scale degree patterns
Professional musicians	——>	Flexible system (interval <—> scale degree)

cises toward each type of listener to develop the other kind of skill. The remainder of Lake's article outlines what appear to be well-designed and specific exercises to develop these skills.

In addition to these exercises, the effective theory teacher will remember the power that "chunking" by harmonic function has for improving students' tonal memories. This skill may be one of the primary factors that differentiates musicians' performance from nonmusicians' in discrimination tasks involving memory for melodies. Classroom exercises designed around arpeggiations of triads and seventh chords in functional harmonic progressions reinforce scale-degree representations; they explicitly teach students to hear the underlying function of melodic arpeggiations, and teach students to group pitches by harmonic function in memory. These exercises may take the form of "echoes" sung by the class after hearing patterns sung by the instructor, then may progress to vocal improvisations by students upon familiar harmonic progressions. Students who improvise melodic patterns from harmonic progressions—vocally or on their instruments—are predisposed to hear melodies given in dictation as elaborations of common harmonic progressions.

Finally, teachers who regularly give melodic dictation should consider the strategies described by Gary Karpinski (#14). In his excellent and detailed exploration of music perception and its implications for melodic dictation, Karpinski cites a number of the studies already discussed here, as well as a few others, to support his step-by-step pedagogy for melodic dictation. In addition to drawing upon psychological experimentation in developing dictation strategies, he puts to use some of the methodology of the sciences, by using the "case study" method to open his discussion of individual differences among students. Each case study subject has a distinctly different place where problems occur in the cognitive chain-of-events that occurs during melodic dictation. Like Lake, Karpinski recommends a series of pretests for each student to identify where in this chain the problems lie; then he recommends specific dictation steps and exercises for remediation that are geared to each problem.

Tonal Hierarchy and Closure

Music theorists have responded in several ways to the burgeoning number of cognitive experiments using musical auditory stimuli. One response—that just illustrated by the publications of Karpinski, Lake, and Larson—is to use this research to confirm one's own intuitions about musical structure and perception, and then to build upon firm cognitive ground a pedagogy that teaches these principles systematically. A second response is to use one's musical intuitions and expert knowledge of musical structure to serve as a critical "outside reader" for the psychologists—to question each assumption, each experimental design, and each conclusion from the musician's point of view. One area of cognitive research that has sparked much heated debate between music theorists and psychologists is the experimental study of perceived tonal hierarchy and closure. It is this body of work to which we turn in part three of the bibliography.

David Butler and Helen Brown's essay, "Describing the Mental Representation of Tonality in Music" (#27), brings together in one place a number of criticisms they have leveled elsewhere regarding the generally impoverished conception of tonality held by some music-psychologists—in particular their understanding of the scale as representative of tonality and their generally limited understanding of the contextual relationship that chromatic pitches have to the diatonic collection. On the subject of scale and tonality, for example, they note that:

> Scales are theorists' attempts to describe the pitch materials in music that already exists. The scale is a handy index of the pitch contents of some piece or pieces of music, but the scale's simplicity comes at a hefty price. The scale is a highly simplistic construction that fails to describe much of what is most important about pitch relations in tonal music.[24]

These statements are reminiscent of Butler's published criticisms (#23, 24) of the "tone probe" technique, used by psychologist Carol Krumhansl to test the perceptual validity of a tonal hierarchy. Figure 5 demonstrates the experimental method and results of her early tone probe studies.[25] Listeners heard a major or minor scale to establish a tonal context, after which a single tone was sounded; subjects were asked to rate how well the tone "fit in" to the tonal context previously heard. Multidimensional scaling techniques were used to create a hierarchy of pitch classes based on perceived distances from tonic. Butler and Krumhansl have engaged in a heated exchange of viewpoints over this issue and others. Not only does a scale fail to approximate a true "tonal context," according to Butler (#23), but there are also effects other than the perception of a tonal hierarchy that might influence listeners' judgments, thus weakening Krumhansl's conclusions. One such possible effect, for example, is the recency effect, which might cause subjects simply to prefer the note most recently heard.

Butler and Brown (#27) also report on their own research into key identification, based on Richmond Browne's (1981) discussion of the tonal implications of the diatonic set.[26] The authors have shown that listeners are better able to determine the implied tonic

---

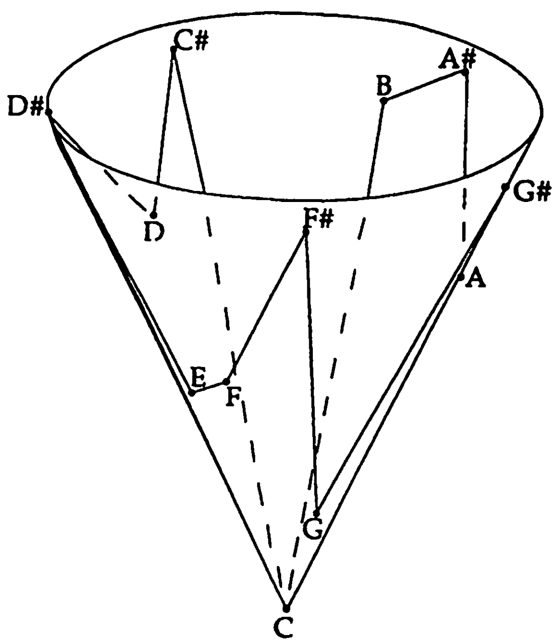
[24]David Butler and Helen Brown, "Describing the Mental Representation of Tonality in Music," in *Musical Perceptions*, ed. Rita Aiello with John Sloboda (New York: Oxford University Press, 1994), 195.

[25]The musical notation of Figure 5 is drawn from David Butler's "Response to Carol Krumhansl," *Music Perception* 7 (1990): 325. The conical representation is reproduced from his "A Study of Event Hierarchies in Tonal and Post-Tonal Music," *Psychology of Music* 18 (1990): 5.

[26]Richmond Browne, "Tonal Implications of the Diatonic Set," *In Theory Only* 5: 3-21.

Figure 5. Krumhansl's "tone probe" technique and tonal hierarchy

c. 1990 by the Regents of the University of California
Reprinted from *Music Perception* 7 (1990) p. 325 by permission



Reproduced from *Psychology of Music* 18 (1990) p. 5
by permission of the editor

of a three-note set of pitches if it contains one of the rare intervals of the diatonic set (the tritone or the half step). Listeners can make this judgment rapidly and accurately; very few tones need to be present in the aural cue, so long as one of the rare intervals is represented. We can intuit the strength of this claim from the power of scale degrees 7 and 4 to predict resolution to 1 and 3, respectively. Butler and Brown also discuss the importance of pitch order in defining a key unambiguously. Brown's experimental research has shown, for example, that two different orderings of an identical pitch collection can suggest two different keys to listeners; a third reordering can be tonally ambiguous. Thus it is not only the presence of a diatonic collection in an experimental stimulus that determines listeners' perception of tonic, but also the order in which the tones of the collection are heard. This may suggest yet another reason to be wary of the randomly generated melodic dictation exercises seen in some computer-assisted instruction programs. Melodies generated by what Dowling calls "a random walk" on the diatonic scale[27] can be highly ambiguous tonally, even when they begin and end on the tonic pitch, since these random processes generally do not take into account the underlying implied harmonic progressions that are part of tonal melodic composition.

Butler and Brown close their article with a discussion of chromaticism.[28] The authors note that chromaticism has been equated with randomness or atonality in some experimental studies. Chromaticism in music, however, generally arises from specific compositional techniques: tonicization of other scale degrees, and mixture of the major and minor modes. Further, the so-called atonal stimuli in some experiments can, in fact, be heard tonally by experienced listeners. The process of chromaticism in tonal music, ac-

---

[27]This is the method whereby diatonic stimuli were generated in Dowling's 1978 study, "Scale and Contour: Two Components of a Theory of Memory for Melodies" (note 10), though the interval sizes were also constrained in Dowling's study.

[28]These ideas are developed more fully in Helen Brown's "Cognitive Interpretations of Functional Chromaticism in Tonal Music," in *Cognitive Bases of Musical Communication*, ed. Mari Riess Jones and Susan Holleran (Washington, DC: American Psychological Association, 1992), 139-160.

cording to these authors, "involves much more than simply adding nondiatonic pitches to . . . dilute the strength of the prevailing tonality. Chromatic chords can lead skilled listeners to expect that some chords will form more appropriate resolutions than others and can thus be clear signals both of the key level and of unfolding harmonic events at the local level."[29]

Nicholas Cook's article on music and psychology (#28) is likewise an overview of various points of contention between theorists and psychologists, including the Krumhansl/Butler polemic just discussed. Cook addresses important issues impacting theories of musical structure as well. For example, he describes an experiment of his own (to which we shall return shortly) that calls into question the perceptual validity of hierarchical structures and the power of tonal closure; he discusses these issues in light of Lerdahl and Jackendoff's and Schenker's theories of tonal music.[30]

Cook also makes an important point about music listening that is often overlooked: musically educated listeners to some extent can choose to listen grammatically or can choose not to do so. This issue of cognitive choice has been addressed experimentally by Rita Aiello and her colleagues, who concluded that

> a high level of music training does not lead to a uniform way of perceiving music. Music education and training broaden students' base of musical knowledge . . . but musicians can use this knowledge selectively, according to the particular musical elements they choose to attend to. In listening to music, musicians focus on different musical elements and perceive the complexity of music according to individual cognitive strategies.[31]

---

[29]Butler and Brown, 208.

[30]This experiment is published as "The Perception of Large-Scale Tonal Closure," *Music Perception* 5: 197-205. Cook discusses the results in light of Fred Lerdahl and Ray Jackendoff's *A Generative Theory of Tonal Music* (Cambridge, MA: MIT Press, 1983) and Heinrich Schenker's *Free Composition*, trans. and ed. by Ernst Oster (New York: Longman, 1979).

[31]Rita Aiello, J.S. Tanaka, Wayne C. Winborne, "Listening to Mozart: Perceptual Differences Among Musicians," *Journal of Music Theory Pedagogy* 4 (1990): 290-291.

Teachers, of course, should be aware of the potential for students to use different strategies (as Lake has pointed out); but in cases where a particular strategy is clearly more effective, teachers need to guide students toward that strategy. Further, because instructors are not always aware of which strategies students are using, they may wish to guide dictation practice in class to ensure that students are learning a full repertoire of strategies (e.g., "on this hearing, determine the meter and first and last scale degrees"). While choosing not to listen grammatically is an acceptable choice in some musical settings, our students need to be taught the cognitive strategies that will enable them to choose to listen grammatically in other settings.

We have discussed two possible ways in which music theorists might respond to music-cognitive research. A third way to respond is to design and run experimental research ourselves. This involves either some training outside our discipline—in experimental design and methods of data analysis, for example—or collaboration with colleagues in psychology. Music-cognitive experimentation enables music theorists to explore directly issues that arise in their aural skills, harmony, or analysis classes. While there seems to be a natural connection between empirical studies of music listening and professional instruction in music listening, the connection between empirical research and "written" theory instruction seems less obvious. Yet several experiments published in recent years pose important questions for the teaching of tonal harmony and analysis. The first of these, by Burton Rosner and Eugene Narmour (#21), investigated the relation between theories of harmonic closure and the perception of closure, citing in their discussion differences of opinion among music theorists over the years regarding such factors as the relative strength of plagal and authentic cadences, as well as the relative strength of scale degrees 7-8 versus 2-1. They trace these issues through the theories of Rameau and Riemann, Schoenberg and Schenker. In their experiment, nonmusician subjects heard two pairs of chords (each pair in a different key) and were asked to judge whether progression A or B sounded more closed. Among their results was the finding that plagal cadences were judged much weaker than authentic ones; subjects even preferred the progression III-I over IV-I for closure if the soprano voice leading was scale degrees 7-8. Melodic closure of 2-1 was not statistically different from 7-8, however. The authors conclude that "the message for Schenkerian theory seems to be that V-I as a generative harmonic constant is empirically tenable, whereas 2-1 as a melodic one . . . is not."[32]

The "message for Schenkerian theory" is not so positive from others engaging these issues. Karno and Konecni (#19) investigated the extent to which the rearrangement of formal sections in the Mozart G Minor Symphony, K. 550 would displease listeners unfamiliar with the work. These rearrangements varied in their nonconformity to tonal principles of structure: among the worst, a version that began with theme 1 from the recapitulation, followed by theme 1 and the closing theme from the exposition, then the development, the recapitulation's transition and closing theme, followed by the exposition's transition, then theme 2 from the recapitulation, ending with theme 2 from the exposition. This version, of course, would end in a different key from that in which it began.

Subjects rated each of five versions along three bipolar scales: pleasing/not pleasing, wish to own/do not wish to own, and interesting/not interesting. The authors reported that in two administrations of the experiment—one with nonmusicians and one with musicians—only when the symphony's original version was heard first was it preferred. When asked to rank the best version overall, the original was slightly preferred, but not by a statistically significant margin. Further, other odd contenders were rated almost as highly. The authors conclude that "it would seem that many (armchair-based) claims made by musicologists about the role of structure in a work's impact can be called into question by empirical research."[33] The authors acknowledge that theories of structure may be less usefully viewed as perceivable musical elements and more usefully seen as conceptual tools of composition. One might question the relationship between the task and the authors' conclusions, however, since they did not direct their subjects to attend to the

---

[32]Burton S. Rosner and Eugene Narmour, "Harmonic Closure: Music Theory and Perception," *Music Perception* 9 (1992): 407.

[33]Mitchell Karno and Vladimir J. Konecni, "The Effects of Structural Intervention in the First Movement of Mozart's Symphony in G Minor, K. 550 on Aesthetic Preference," *Music Perception* 10 (1992): 72.

music's structural features in the experiment; that is, they did not ask subjects to rate the succession of melodies, their perception of narrative drama, or even their sense of completion. Rather, they focused on vague aesthetic notions of pleasingness and interest.

It is a music theorist, Nicholas Cook (#18), who explicitly asked listeners to judge tonal closure, among other elements, in short compositions from one to six minutes in duration. Noting the importance of monotonality and large-scale tonal closure as fundamental principles underlying most theories of tonal music, Cook designed an experiment using short compositions by Beethoven, Brahms, Haydn, and other common-practice composers, which were performed for subjects in two versions: the original and an altered version in which the second half of the work was transposed. Thus the altered compositions ended in keys different from those in which they began. Nineteen music students were played each of the composition pairs and asked to state which of the two versions they preferred in each of the following categories: pleasure, expressiveness, coherence, and sense of completion. Only for the very shortest of the compositions was there any significant link between tonal closure and aesthetic response. Cook asks the reader: "If the principle of tonal closure has little or no perceptual validity at the larger time scales found in most tonal compositions, is there not something radically wrong with a theory that ascribes fundamental aesthetic importance to it?"[34] Perhaps we should question whether Cook's question is really fair: do theories of tonal structure really intend to model aesthetic response? Cook's data analysis does not show that listeners are unable to discern that a composition has ended in a different key—only that listeners do not base their aesthetic response upon this criterion.

In his article's concluding discussion, Cook also raises this point, and notes that theories of musical structure need not be theories of musical perception as well. Theories of tonal structure are not scientific in intent, in Cook's opinion; rather, they are pedagogical, since it is essential that composers and performers alike have an abstract conception of a work's structure in order to control the composition's

---

[34]Nicholas Cook, "The Perception of Large-Scale Tonal Closure," *Music Perception* 5 (1987): 203.

large-scale dynamic and temporal processes. Cook concludes by noting in connection with sonata form that

> the disposition of textures and thematic materials, the patterning of loud and soft passages and of high and low tessitura, the pacing of tension and relaxation—all of these aspects of a sonata are organized around the tonal plan and serve to project its structural closure in a directly perceptible manner. Hence if large-scale tonal relations are not in themselves audible, that does not necessarily mean that they are of no musical significance: it may just be that their influence on what is heard is an indirect one.[35]

Cook's empirical research into the perceptual validity of tonal closure does not invalidate what we do in harmony and analysis classrooms. As he notes, large-scale form is articulated by many perceivable features; these should be just as much a part of aural analysis as determination of the key of the second theme group. Further, since the students in our classrooms are primarily performers, our detailed examination of motivic unity and development, harmonic areas tonicized, phrase structure, and large-scale formal relations will help provide these young performers with a framework for memorization and with important conceptual information for musical interpretation.

Studies on Individual Differences

One final way in which music-cognitive research may help inform pedagogical method is in the study of individual differences among students. Within most music theory classrooms, there are distinct subgroups of students who share particular characteristics. One such characteristic is absolute pitch (AP) recognition. How should instructors handle dictation training for these students? At some institutions, these students are given a placement test and exempted from aural skills training altogether, because of their ability to sing at sight and to take accurate dictation. Yet if we ask AP students to listen to a harmonic progression and write Roman numerals only—without benefit of staff paper—they are often com-

[35]Ibid., 204.

pletely unable to do this, because they have not been trained to hear harmonic function. These students tend to transcribe letter names as quickly as possible, then return to their transcriptions and analyze for Roman numeral function. Even in the simplest of tasks, like interval identification, they tend to write (or think) letter names first and then analyze for the correct label. Some AP students cannot complete dictation tasks if a turntable or tape machine plays slightly fast or slow, resulting in a pitch level that is slightly sharp or flat from the key announced by the instructor.

Teachers of sight singing and dictation who have experienced these problems with AP students will be interested to know of an article entitled "Absolute Pitch as an Inability: Identification of Musical Intervals in a Tonal Context" (#35), which confirmed in an experimental setting exactly what we have observed in our classrooms. The experimenter, Ken'ichi Miyazaki, asked AP listeners to name melodic intervals in three key contexts: C Major, F# Major, and an entirely quarter-step-flat E Major. Each interval was preceded by a harmonic progression to establish one of these keys, and the first note of the interval was scale degree 1 of that key. Subjects were asked to name the interval, given only the choices minor third, major third, or perfect fourth.[36] The catch here is that the intervals heard were mistuned by 20-cent increments (a half step contains 100 cents).[37] Miyazaki scored an interval identification as correct if the interval heard was 40 cents wider or narrower than the interval named. Since an equal-tempered major third contains 400 cents, "major third" as a response in this experiment was judged correct if the interval played spanned anywhere from 360 to 440 cents.

---

[36]Miyazaki notes one possible confounding variable in the design of his experiment: subjects were given the option of responding with interval names or with movable-do solfège names. However, many of the subjects were trained in fixed-do solfège and may not have understood how to use the movable syllables in this context.

[37]This type of experimental method has many precedents in perception research; it tests "categorical perception": the ability of perceivers to classify something that occurs along a continuum into discrete categories or classes. For example, a continuum of microtonal intervals exists between C and its octave above, but we tend to categorize these intervals perceptually into eleven discrete types.

Miyazaki found that AP subjects scored significantly lower and took significantly longer to respond in F# Major and the quarter-tone-flat E Major than in C Major. Further, the non-AP control group showed a considerably higher level of performance than the AP group across all key contexts. Some of the AP subjects showed considerable variability in their responses; their inconsistency was comparable to that of subjects in a related study who had had no musical training at all. The author hypothesizes that even though this task was designed to engage a relative pitch strategy, some AP listeners were unable to make use of that strategy. Miyazaki suggests that all AP listeners have some degree of relative pitch, but that this skill is less highly developed (or actually suppressed) in some listeners.

One music theorist doing collaborative research in a medical center has come to the same conclusion via considerably different experimental means. Ed Hantz and his collaborators (#32) discovered that absolute pitch listeners activate differently shaped brain wave forms (the P3 wave) in an interval identification task, or they do not activate this short-term "working memory" wave form at all. They hypothesize that AP listeners instead retrieve pitch names from long-term memory, then calculate interval size. This use of an AP strategy, even in the face of a relative pitch task such as interval naming, likewise suggests that some AP listeners are unable to choose a relative pitch strategy.

What are we to make of Miyazaki's finding that AP listeners were less well able to identify intervals from an F# tonic than from a C tonic? There is now a body of experimental evidence showing that AP subjects identify white-key pitches more quickly and accurately than black-key. Annie Takeuchi and Stewart Hulse (#36) hypothesize that there exists a window of opportunity in childhood—similar to the "critical period" for language acquisition—during which time absolute pitch recognition is acquired. Since children typically begin music instruction using the white keys of the piano and reading music with few accidentals, these are the pitches first acquired; it may be that the window of opportunity closes before pitch memory for the black keys is acquired. Older AP listeners may determine black-key note names by half-step displacement from the white-key notes, thus accounting for the slower recognition times. Miyazaki concurs with the early music training hypothesis and notes that childhood AP acquisition probably results in suppressing the development of relative pitch. In fact, he goes so far as to suggest that children who have begun music lessons from an early age should be given explicit training in relative pitch—a "remedial treatment for children having AP."[38]

In what other sense might absolute pitch be regarded as an "inability"? Recall my experiment on recognition of invariant-contour melodies described previously (#34); subjects heard a tonal or atonal melody, followed by either an exact transposition or altered transposition, then were asked to make a same/different discrimination. AP subjects scored significantly higher than non-AP subjects in discriminating transposed tonal melodies from same-contour lures. In the nontonal condition, however, this advantage completely disappeared. In fact, there was no significant difference between AP musicians and nonmusicians' performance in the nontonal condition. Like Miyazaki, I conclude that AP listeners are able to use a combination of absolute pitch and relative pitch strategies, depending upon the task. The data from this experiment suggest that AP musicians use pitch-class and scale-degree information primarily, with less reliance upon memory of interval successions. Upon hearing a transposed tonal comparison melody, AP subjects might choose to transpose the original melody's remembered pitch classes to a new key or—more likely—they might shift to a scale-degree recognition strategy, which is invariant across the key change. For the nontonal melodies, however, AP subjects may have found themselves unable to transpose quickly and accurately enough to discriminate transposed targets from lures. Because of the nontonal context, they were unable to use a scale-degree strategy, and the results suggest that they were less able to use an interval strategy. I conclude that at least some of the AP subjects were unable to use a flexible system that incorporated interval information, since their ability to discriminate declined so markedly in the atonal condition.[39]

---

[38]Miyazaki, 70.

[39]Elizabeth West Marvin, "Cognitive Strategies for Recognition of Invariant-Contour Melodies: Tonal Context Affects Performance Among

What implications do these studies have for theory pedagogy? In a paper presented at the Second International Conference for Music Perception and Cognition, Philip Baczewski and Rosemary Killam (#30) reported on the "extreme and negative reactions" of their AP subjects who were instructed during an experiment to notate a two-voice dictation in a key other than that in which it was played.[40] Clearly, these musicians had never been asked in a classroom situation to think and write in other keys. The preponderance of data reported in AP experiments suggests that few absolute pitch subjects have been explicitly trained in relative pitch listening strategies. These musicians simply rely on their AP, whether the AP seems to be "working" for them or not. Research suggests that AP musicians with under-developed relative pitch skills will encounter cognitive difficulties when performing at Baroque pitch, for example, or on non-equally tempered instruments; they may struggle when asked to transpose from score, and have difficulty hearing transposed motivic relationships in nontonal music.

Responsible aural skills teachers need to address this deficiency by incorporating into their lesson plans specific exercises to help AP listeners develop their relative pitch skills. For example, many classroom exercises in the aural skills classes can be designed to promote maximum musical interaction between teacher and student without the intrusion of any staff notation at all. Tonal echo patterns and vocal improvisation exercises upon simple harmonic progressions can be sung in a variety of keys on scale-degree numbers or (movable-do) solfège syllables without notation. Exercises in which the instructor sings chordal arpeggiations on a neutral syllable, asking students to echo back by singing on scale-degree numbers or solfège syllables, encourage students to map these patterns onto harmonic function rather than note names. These are good preliminary exercises to stretch AP students' relative pitch

---

Musicians and Non-Musicians," paper presented at the annual meeting of the Society for Music Perception and Cognition, Philadelphia, PA, 1993.

[40]Philip Baczewski and Rosemary N. Killam, "Perception of a Musical Construct by Musicians with Absolute Pitch," paper presented at the Second International Conference on Music Perception and Cognition, Los Angeles, CA, 1992.

skills. In addition, from the first day of the freshman year, students should regularly be asked to sing melodies at sight in keys other than notated. This technique makes no difference to relative pitch students, but it begins to stretch the AP students into new ways of thinking about musical relations while the melodies are fairly simple and diatonic, and it has the added advantage of allowing instructors to move melodies notated in tessituras too high or too low into more comfortable keys for singing.

In dictation classes, melodies may be given without announcing a key at all during much of the freshman year. Students would be expected to provide rhythmic notation and scale-degree numbers only. This dictation strategy views transcription onto the staff as a final extra step. Once the tune's scale-degree structure has been notated numerically, teachers may call out any key at random and ask students to notate the melody in this key. This dictation procedure brings to mind Gary Karpinski's article (#14), discussed earlier.[41] His step-by-step dictation pedagogy is excellent for AP students as well as for remedial students, because it explicitly exercises relative pitch relations. Finally, as instructors choose keys for singing and playing dictation, the white-key/black-key finding ought to be kept in mind; music-cognitive research suggests that black-key keys may be more effective in forcing AP students to rely upon relative pitch, though these keys may be more difficult at first.

Before leaving the subject of individual differences among students, let us turn to the question of sex differences. Are there possible differences between the ways men and women perform mu-

---

[41]Another interesting parallel with the Karpinski research may be found in Daniel Levitin's "Absolute Memory for Musical Pitch: Evidence from the Production of Learned Melodies," *Perception & Psychophysics* 56 (1994): 414-423. Karpinski argues that one distinct type of dictation problem is the ability to sing back a melody accurately, but the inability to associate scale-degree numbers or solfège syllables with the melody just sung—that is, the inability to assign to each pitch a meaningful name. Levitin shows that even nonmusicians have the ability to retain melodies *at pitch*, which he demonstrates by having subjects sing pop music songs they know well. Even subjects with no musical training often sing these tunes in the correct key; they have the ability to remember a pitch and key, but they do not have the training (or perhaps the ability) to assign a meaningful name to that pitch.

sic-cognitive tasks? Indeed, there is a body of research in sex differences outside of our field. Doreen Kimura's 1992 article in *Scientific American* (#39) provides a concise summary of this research. Figure 6 summarizes the cognitive tasks in which sex differences have been shown. In related research, Thomas Bever (#37) demonstrated that in spatial navigational tasks women and men perform equally well, but that they use different mental strategies to keep track of their positions. After navigating a life-size maze, men were better able to identify a map of the route just traversed; women, on the other hand, were better able to place in the correct order a series of pictures showing items passed along the way. Bever concluded that women more often navigate through a maze by means of remembered landmarks, while men use global maps constructed in memory by means of directional vectors. Are there any implications for music listening here? Can memory for music be considered a spatial task that might exhibit sex differences in strategy?

A few links between spatial perception and music study have been found previously. For example, Marianne Hassler and her colleagues (#38) have undertaken a longitudinal study of young children through puberty; these researchers have found significant correlations between musical talent and high visual-spatial ability as measured by standardized tests. Interestingly, the correlation is higher among children who compose or improvise than among other young performers. No sex differences were found in children before puberty, but after puberty spatial visualization was stronger in females than males, leading to a sex difference in favor of the females. The issue of whether music study by children somehow improves spatial cognitive functioning is the subject of an on-going series of experiments by Frances H. Rauscher and her colleagues at the University of California at Irvine.[42] This group of researchers studied three-year-olds, two-thirds of whom were given eight months of music training (group singing and private keyboard lessons). When tested on a spatial reasoning task that required them

[42]Rauscher's research has been reported widely in the popular press; see, for example, "Brain Music: Listening to Intricate Mozart Melodies Can Improve Ability to Learn" by Richard A. Knox, which appeared in the Rochester, NY *Democrat and Chronicle* (Monday, September 12, 1994) as a reprint of a *Boston Globe* story.

Figure 6. Cognitive tasks exhibiting sex differences (Kimura, 1992)

Tasks favoring women's performance:

- tests of perceptual speed (such as matching items)
- ideational and verbal fluency (for instance, listing words beginning with the same letter)
- precision manual tasks (placing pegs in holes)
- mathematical calculations

Tasks favoring men's performance:

- spatial tasks (such as mentally rotating an object)
- target-directed motor skills
- disembedding tasks (finding a simple shape in a complex form)
- mathematical reasoning

to assemble picture puzzles, the test scores of children engaged in music training had dramatically improved after eight months, while the other children showed no significant improvement. The researchers hypothesize that spatial tasks like the assembling of puzzles are similar to music performance, since both require "forming an ideal mental representation of something which is eventually realized."[43]

To explore the issue of spatial representation of music in memory among college-aged adults, I ran an experiment (#40) designed to engage listeners in a spatial task using musical stimuli. As in previous experiments, listeners heard an unfamiliar seven-note tonal melody, immediately followed by a comparison melody. Figure 7 gives a sample standard melody followed by four possible comparison melodies. The task was to identify on an answer sheet whether the comparison melody was an exact transposition, called "same," or whether it was a spatial transformation: backwards, inverse, or backwards-inverse presentation (in other words: retro-

[43]Ibid.

grade, inversion, or retrograde inversion). Figure 8 presents the results, ranked from easy to difficult for male and female musicians. Musicians were able to discriminate in all four conditions at above chance levels. There were no significant sex differences in the T- and I- conditions. There was, however, a significant difference in the R-condition, with males better able to discriminate; there was a strong, though not statistically significant, trend in favor of males' discrimination of the RI-transformation as well. To summarize, then, significant differences were found between the sexes, but only for the transformations that reversed the pitches' temporal order.

Subjects were also asked to complete a questionnaire; among the questions asked was a strategy question. Subjects' responses were divided into three categories: memorization of distinct "segments" (a landmark strategy), visualization of the entire "contour" or shape of the melody (a vector strategy), and "other." This categorization revealed more flexibility in strategy among males; over

Figure 7. Melodic transformations from Marvin (1993)

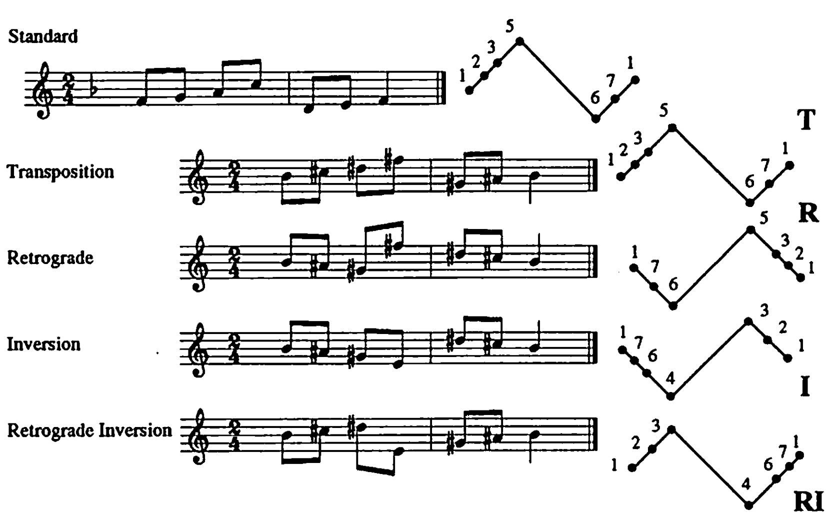


Figure 8. Marvin's ranking of transformations
(easy to difficult)

Female subjects: T —> I —> R <—> RI

(no significant difference between R/RI)

d′ means:* 1.95 .97 .34 .31

($p<.0005$, $p<.0001$)

Male subjects: T —> R <—>I <—> RI

(no significant difference between R/I or I/RI)

d′ means:* 2.3 .90 .85 .62

($p<.0001$)

*D-prime values control for subjects' possible guessing strategies by plotting correct identifications (hits) against false positives (false alarms). They range in value from -3.29 to +5.73, where 0 represents no discrimination (chance). D′ values used here were derived from the conversion tables in M. J. Hacker & R. Ratcliff, *Perception & Psychophysics* 26 (1979): 168-170.

half reported a segment strategy, one-third a vector strategy, and the remainder "other." Two-thirds of the women, on the other hand, reported a segment strategy, with the remaining third divided between "contour" and "other" strategies. Statistical analysis of these data suggests that males and females generally do use different cognitive strategies to complete this task.

The poorer discrimination rates by females in this experiment can be attributed, in part, to the fact that a segment (landmark) strategy is a less effective strategy in this task than vector representation (contour), since some segments remain invariant under retrograde and retrograde inversion. In Figure 7, for example, the first melodic shape would be represented in scale degrees as 1-2-3-5-6-7-1 and the retrograde inversion as 1-2-3-4-6-7-1; note that they both begin with scale degrees 1-2-3, and end with 6-7-1. If those were the segments that the listener chose to remember, confusions between same and backwards-inverse would result. The same problem arises in the backwards and inverse conditions, with 1-7-6-5-3-2-1 and 1-7-6-4-3-2-1. Further, a melody like 1-2-7-5-4-2-1—which has 1-2 at the beginning and 2-1 at the end—would result in errors in same/retrograde discrimination if the listener chose the beginning/ending dyads to remember as landmarks. These two factors, then, may account for women's lower scores in the R- and RI-conditions, since the female subject group more frequently reported using the landmark strategy. The more effective strategy for this task takes into account the contour of the whole, not just the ups and downs, but also the relative distance up and down: a vector representation.

If men and women are cognitively predisposed to perform spatial transformations on remembered melodies in fundamentally different ways, then teachers of aural skills need to broaden their methodologies to embrace both ways of thinking about (and hearing) such transformations—either as global maps or as more localized landmarks. In the task of identifying the new key in a piece that has modulated, for example, the teacher who explains the strategy only in the global terms of memorizing the initial key, retaining it in memory, and then comparing it to the new key may be teaching with a male bias. Presenting also an alternative landmark-based strategy that focuses on reinterpretation of scale degrees during the modulatory shift might be more inclusive. In harmonic dictation, a global strategy might focus first on large harmonic goals, while a landmark strategy might take note of chord qualities, altered tones, or nonharmonic tones along the way. These examples are not intended to imply that there is unequivocally one "female" way to hear a relationship and another "male" way; rather, it is to assert that different classes of listeners may process music differently (absolute pitch listeners may use still another strategy). The effective teacher should present several alternative methods.

Let us close by returning to Cook's essay on music and psychology (#28) because it draws important distinctions between what psychology and music theory should be and should do. Cook notes that

> musical perception is pluralistic and fluid; listeners make use of multiple cognitive frameworks . . . and shift their strategies from one moment to the next. . . . A psychological model of perception needs to embody these characteristics. But there is no reason why music theory should be like this. . . . The aim of music theory, then, is to go beyond perception.[44]

While theories of musical structure need not model cognitive processes, an awareness of these processes can considerably ease a music theory teacher's task. Good experimental research in music cognition can make music theory instructors aware of the ways in which our students are predisposed to hear and interpret music; it can remind us that different classes of listeners (musicians and nonmusicians, for example) hear and interpret music using entirely different cognitive strategies. The more our pedagogical methods can build upon students' predispositions, yet lead them toward strategies that have been shown most effective for skilled musicians, the better our teaching will be.

[44]Cook, "Music and Psychology," 89.

SELECTED BIBLIOGRAPHY

This listing of suggested readings is organized topically, but articles are numbered consecutively throughout to facilitate their citation in the body of the article.

I. Overviews of the field

(1) Aiello, Rita with Sloboda, John A., Eds. *Musical Perceptions.* New York and Oxford: Oxford University Press, 1994.

(2) Butler, David. *The Musician's Guide to Perception and Cognition.* New York: Schirmer Books, 1992. See also the review by Elizabeth West Marvin, *Journal of Music Theory Pedagogy* 7 (1993): 119-128.

(3) Deutsch, Diana, Ed. *The Psychology of Music.* New York: Academic Press, 1982.

II. Research on melodic perception and recall

A. Experimental studies:

(4) Balzano, Gerald J., and Liesch, Barry W. "The Role of Chroma and Scale Step in the Recognition of Musical Intervals In and Out of Context." *Psychomusicology* 2 (1982): 3-32.

(5) Bartlett, James C., and Dowling, W. Jay. "Recognition of Transposed Melodies: A Key Distance Effect in Developmental Perspective." *Journal of Experimental Psychology: Human Perception and Performance* 6 (1980): 501-515.

(6) Croonen, W. L. M., and Kop, P. F. M. "Tonality, Tonal Scheme, and Contour in Delayed Recognition of Tonal Sequences." *Music Perception* 7 (1989): 49-68.

(7) Cuddy, Lola. "On Hearing Pattern in Melody." *Psychology of Music* 10 (1982): 3-10.

(8) Dowling, W. Jay. "Context Effects on Melody Recognition: Scale-Step versus Interval Representations." *Music Perception* 3 (1986): 281-296.

(9) Dowling, W. Jay. "Melodic Contour in Hearing and Remembering Melodies." In *Musical Perceptions*, 176-190. [See (1) for full citation.] See also Rita Aiello's "Introduction," 173-176.

(10) Killam, Rosemary; Lorton, Paul V. Jr.; and Schubert, Earl D. "Interval Recognition: Identification of Harmonic and Melodic Intervals." *Journal of Music Theory* 19 (1975): 212-234.

(11) Sloboda, John and Parker, D. H. "Immediate Recall of Melodies." In *Musical Structure and Cognition*, ed. Peter Howell, Ian Cross, and Robert West. London: Academic Press, 1985.

(12) Takeuchi, Annie and Hulse, Stewart H. "Key-Distance Effects in Melody Recognition Reexamined." *Music Perception* 10 (1992): 1-24.

B. Implications for the theory classroom

(13) Butler, David and Lochstampfor, Mark. "Bridges Unbuilt: Aural Training and Cognitive Science." *Indiana Theory Review* 14 (1993): 1-18.

(14) Karpinski, Gary. "A Model for Music Perception and its Implications in Melodic Dictation." *Journal of Music Theory Pedagogy* 4 (1990): 191-229.

(15) Lake, William E. "Interval and Scale-Degree Strategies in Melodic Perception." *Journal of Music Theory Pedagogy* 7 (1993): 55-67.

(16) Larson, Steve. "Scale-Degree Function: A Theory of Expressive Meaning and its Application to Aural Skills Pedagogy." *Journal of Music Theory Pedagogy* 7 (1993): 69-84.

III. Tonal hierarchy and closure

A. Experimental studies

(17) Butler, David. "A Study of Event Hierarchies in Tonal and Post-Tonal Music." *Psychology of Music* 18 (1990): 4-17.

(18) Cook, Nicholas. "The Perception of Large-Scale Tonal Closure." *Music Perception* 5 (1987): 197-206.

(19) Karno, Mitchell and Konecni, Vladimir J. "The Effects of Structural Intervention in the First Movement of Mozart's Symphony in G Minor, K. 550 on Aesthetic Preference." *Music Perception* 10 (1992): 63-72.
(20) Krumhansl, Carol. *Cognitive Foundations of Musical Pitch.* New York: Oxford University Press, 1990.
(21) Rosner, Burton S. and Narmour, Eugene. "Harmonic Closure: Music Theory and Perception." *Music Perception* 9 (1992): 383-412.
(22) Serafine, Mary Louise; Glassman, Noah; and Overbeeke, Cornell. "The Cognitive Reality of Hierarchic Structure in Music." *Music Perception* 6 (1989): 397-430.

B. Krumhansl versus Butler polemic

(23) 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 (1989): 219-242.
(24) Butler, David. "Response to Carol Krumhansl." *Music Perception* 7 (1990): 325-338.
(25) Krumhansl, Carol. "Tonal Hierarchies and Rare Intervals in Music Cognition." *Music Perception* 7 (1990): 309-324.

C. Critiques by music theorists

(26) Brown, Helen. "Cognitive Interpretations of Functional Chromaticism in Tonal Music." In *Cognitive Bases of Musical Communication*, ed. Mari Riess Jones and Susan Holleran, 139-160. Washington, D.C.: American Psychological Association, 1992.
(27) Butler, David and Brown, Helen. "Describing the Mental Representation of Tonality in Music." In *Musical Perceptions*, 193-212. [See (1) for full citation.]
(28) Cook, Nicholas. "Music and Psychology: A Mutual Regard?" In *Musical Perceptions*, 66-95. [See (1) for full citation.]
(29) Hantz, Edwin. "Studies in Musical Cognition: Comments from a Music Theorist." *Music Perception* 2 (1984): 237-244.

IV. Studies on individual differences

A. Absolute Pitch

(30) Baczewski, Philip and Killam, Rosemary N. "Perception of a Musical Construct by Musicians with Absolute Pitch." Paper presented at the Second International Conference on Music Perception and Cognition, Los Angeles, CA, 1992.

(31) Benguerel, André-Pierre, and Westdal, Carol. "Absolute Pitch and the Perception of Sequential Musical Intervals." *Music Perception* 9 (1991): 105-120.

(32) Hantz, Edwin; Crummer, Gary; Wayman, John; Walton, Joseph; and Frisina, Robert. "Effects of Musical Training and Absolute Pitch on the Neural Processing of Melodic Intervals: A P3 Event-Related Potential Study," *Music Perception* 10 (1992): 25-42.

(33) Levitin, Daniel. "Absolute Memory for Musical Pitch: Evidence from the Production of Learned Melodies." *Perception & Psychophysics* 56 (1994): 414-423.

(34) Marvin, Elizabeth West. "Cognitive Strategies for Recognition of Invariant-Contour Melodies: Tonal Context Affects Performance Among Musicians and Non-Musicians." Paper presented at the annual meeting of the Society for Music Perception and Cognition, Philadelphia, PA, 1993.

(35) Miyazaki, Ken'ichi. "Absolute Pitch as an Inability: Identification of Musical Intervals in a Tonal Context." *Music Perception* 11 (1993): 55-72.

(36) Takeushi, Annie and Hulse, Stewart. "Absolute-Pitch Judgments of Black-and White-Key Pitches." *Music Perception* 9 (1991): 27-46.

B. Sex Differences

(37) Bever, Thomas G. "The Logical and Extrinsic Sources of Modularity." In *Modularity and Constraints in Language and Cognition*, ed. M. Gunnar and M. Maratsos, 179-211. Hillsdale, NJ: Lawrence Erlbaum and Associates, 1992.

(38) Hassler, Marianne; Birbaumer, N., and Feil, A. "Musical Talent and Visual-Spatial Ability: Onset of Puberty." *Psychology of Music* 15 (1987): 141-151.

(39) Kimura, Doreen. "Sex Differences in the Brain." *Scientific American* 267 (1992): 119-125.

(40) Marvin, Elizabeth West. "The Effect of Sex Differences on Recognition of Melodic Transformations." *Proceedings* of the European Society for the Cognitive Sciences of Music, Third International Music Cognition and Perception Conference, Liège, Belgium (1994): 449-450.