



UNIVERSITY OF CALIFORNIA PRESS  
JOURNALS + DIGITAL PUBLISHING

---

Musical Features That Facilitate Melody Identification: How Do You Know It's "Your" Song When They Finally Play It?

Author(s): MATTHEW D. SCHULKIND, RACHEL J. POSNER and DAVID C. RUBIN

Source: *Music Perception: An Interdisciplinary Journal*, Vol. 21, No. 2 (Winter 2003), pp. 217-249

Published by: [University of California Press](#)

Stable URL: <http://www.jstor.org/stable/10.1525/mp.2003.21.2.217>

Accessed: 19/05/2015 22:45

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



University of California Press is collaborating with JSTOR to digitize, preserve and extend access to *Music Perception: An Interdisciplinary Journal*.

<http://www.jstor.org>

---

## Musical Features That Facilitate Melody Identification: How Do You Know It's "Your" Song When They Finally Play It?

---

MATTHEW D. SCHULKIND  
*Amherst College*

RACHEL J. POSNER & DAVID C. RUBIN  
*Duke University*

What information do listeners use to "Name that Tune"? This question was investigated in a two-phase experiment. In Phase 1, the participants heard familiar melodies that were played on a note-by-note basis until they were identified. In Phase 2, each note of the melody was analyzed along a variety of musical dimensions. Multiple regression analyses determined which musical characteristics predicted identification performance. Identification was most strongly associated with notes located at phrase boundaries, notes that completed alternating sequences of rising and falling pitches, and metrically accented notes. As well, identification peaked after listeners heard moderate amounts of information (i.e., 5–7 notes). The data suggest that melody identification is a holistic, all-or-none process and that parallels can be drawn between melody and spoken word identification. Implications for current theories, future research, and the relationship between music perception and melody identification are discussed.

Received July 21, 2000, accepted August 27, 2003

RESEARCH on object identification has tended to focus on visual and linguistic objects to the exclusion of other types of stimuli such as music. This is surprising given the frequency with which musical stimuli are encountered outside of the laboratory (A. J. Cohen, 1990; Gianetti, 1982; Stewart & Furse, 1986; Yalch, 1991) and the ease with which they are identified inside the laboratory (Andrews, Dowling, Bartlett, & Halpern, 1998; Bartlett, Halpern, & Dowling, 1995; Bartlett & Snelus, 1980; Cuddy,

Address correspondence to Matthew D. Schulkind, Department of Psychology, Amherst College, Amherst, MA 01002. (e-mail: mdschulkind@amherst.edu)

ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

1993; Hébert & Peretz, 1997; Maylor, 1991; Schulkind, 1999; Schulkind, Hennis, & Rubin, 1999; White, 1960). The ease with which melody identification is achieved suggests that listeners retain a great deal of information about the music they know. In fact, research shows that listeners are accurate at reproducing both the pitch and tempo of familiar music (Halpern, 1988, 1989; Levitin, 1994; Levitin & Cook, 1996). However, these data do not explain how listeners use this information to match a physical stimulus with its corresponding mental representation.

Some researchers have addressed this issue by assessing melody identification after selectively altering one or another aspect of familiar melodies (Hébert & Peretz, 1997; Schulkind, 1999; White, 1960). Although these studies demonstrate that both pitch and temporal manipulations influence identification, two aspects of this method limit the generalizability of their conclusions. First, the manipulations used in these experiments were fairly gross. A common manipulation involved preserving the pitch pattern while equating the duration of every note (Hébert & Peretz, 1997; Schulkind, 1999; White, 1960), or preserving the temporal pattern while equating the pitch of every note (Hébert & Peretz, 1997; White, 1960). Second, this approach assumes that manipulating one aspect of the melody (e.g., rhythm) will have negligible effects on the perception of other musical features. Many experiments demonstrate that musical features share interactive rather than additive relationships (Boltz, 1991; Jones & Boltz, 1989; but see also Hébert & Peretz, 1997; Palmer & Krumhansl, 1987a, 1987b). Therefore, the strategies used to identify *altered* versions of familiar melodies may be different from those used to identify *unaltered* versions of the same melodies.

Conducting research using unaltered versions of familiar melodies is complicated because the multilayered, interactive structure of music provides a wide range of musical features that could contribute to identification (Jones, 1987). The current experiment overcame this problem by relying on the vast music perception literature to narrow the focus of the investigation. Simply stated, a musical feature that is difficult to perceive will be unlikely to aid identification. However, this argument does not imply that all easily perceived musical features play a role in melody identification. For example, although listeners can easily perceive transpositions to different keys and/or registers, these changes have little effect on identification (Dowling, 1978). In addition, the experiment was designed to make use of multiple regression analyses, which allow a large number of potential predictor variables to be assessed simultaneously.

The main goals of the current experiment were to answer two questions about melody identification that lie at the heart of object identification research, in general (Tarr & Vuong, 2002). First, what are the fundamental units that are used to identify familiar melodies? Second, how is information in the environment compared with stored representations to allow identification?

### What Are the Fundamental Units for Melody Identification?

Given that music is composed of serially ordered, discrete pitches in time, individual notes would seem to be the logical answer to the fundamental unit question. However, it is possible that listeners do not attend to individual notes, but rather analyze the overall shape or structure of the melody. These contrasting approaches to melody identification roughly correspond to the distinction between analytic and holistic processing, an issue with a long history in studies of both language and music processing (Banich & Heller, 1998; Bever & Chiarello, 1974; Gates & Bradshaw, 1977; Kimura, 1961, 1964). Analytic processing of music would occur if the listener broke the melody down into its constituent components, analyzed each component on a note-by-note basis, and summed the information contained in each note to reach an understanding of the stimulus. For example, listeners might represent “Frosty the Snowman” as a sequence of frequencies or chromas (e.g.,  $F_4 D_4 E\flat_4 F_4 B\flat_5 A_5 B\flat_5 C_5 B\flat_5 A_5 G_4 F_4$ ; for simplicity, information regarding temporal characteristics of the notes is omitted from this discussion). This possibility is dubious given that listeners easily recognize familiar melodies transposed to new keys or registers (Dowling, 1978). Alternatively, listeners may analyze the intervals between successive notes. As such, listeners might represent “Frosty the Snowman” as either a sequence of intervals (m3 m2 M2 P4 m2 m2 M2 M2 m2 M2 M2), a sequence of contour directions ( $\downarrow \uparrow \uparrow \uparrow \downarrow \uparrow \uparrow \downarrow \downarrow \downarrow$ ), or a combination of interval and contour information ( $\downarrow m3 \uparrow m2 \uparrow M2 \uparrow P4 \downarrow m2 \uparrow m2 \uparrow M2 \downarrow M2 \downarrow m2 \downarrow M2 \downarrow M2$ ). Research is consistent with the idea that listeners represent contour and interval information at this level of detail and that the dominant representation varies with the familiarity and length of the melody (Dowling, 1978; Dowling & Bartlett, 1981; Edworthy, 1982, 1985). Thus, an analytic approach to melody identification would consist of matching the string of elements in the physical stimulus with the corresponding sequence stored in memory. For example, if the first three elements of a melody were  $\downarrow m3 \uparrow m2 \uparrow M2$ , “Frosty the Snowman” would be a candidate response. This candidate title would be retained if the next interval was  $\uparrow P4$ , but eliminated if it was anything else.

Although the possibilities described in the preceding paragraph differ in terms of the amount of detail included in the mental representation, they all agree that melodies are stored as ordered strings of discrete, coherent elements (notes, intervals, etc.). However, it is also possible that information about individual notes is lost as they are blended into a unified percept or Gestalt. Consider the following symbol: +. People generally perceive this as a plus sign, rather than as two intersecting line segments at right angles to one another. The individual elements are lost when they are combined to form a meaningful object. Similarly, the individual elements of a melody may be blended to form a larger object that is represented in terms of its

global pitch and temporal structure rather than as a summation of its individual elements. Thus, a holistic approach to melody identification might be achieved by matching information about the global structure of a melody with its corresponding psychological representation.

Research on music cognition generally shows that the way music is processed depends on the abilities of the participants and the demands of the experimental task (Bever & Chiarello, 1974; Gates & Bradshaw, 1977; Robinson, 1977; Robinson & Solomon, 1974). For example, Peretz and Morais (1987) conducted a series of dichotic listening studies similar to the monaural experiments run by Dowling (1978; Dowling & Fujitani, 1971). Peretz and Morais reasoned that whereas remembering the contour of a melody requires holistic processing, remembering the interval structure requires analytic processing. Given that the left and right hemispheres appear to be specialized for analytic and holistic processing, respectively (e.g., Banich & Heller, 1998), Peretz and Morais predicted that a left hemisphere (right ear) advantage should be observed when distinguishing between two melodies that share the same contour, but possess different intervals. Conversely, a right hemisphere (left ear) advantage should be observed when distinguishing between melodies that differ in contour. These predictions were generally supported under several testing conditions. In another set of studies, Peretz, Morais, and Bertelson (1987) instructed listeners to either pay attention to isolated notes, pay attention to the contour, or make aesthetic judgments of the melody; a control group was given no explicit instructions. Whereas the first task would engage analytic processing, the other two tasks would be more likely to involve holistic processing. On a subsequent recognition memory test, the listeners who paid attention to isolated notes (analytic processing) showed the expected right ear/left hemisphere advantage compared with the control group. However, the contour and aesthetic judgment groups did not show the expected left ear/right hemisphere advantage. Peretz et al. argued that these results indicate that right hemisphere, global processing of music is automatic; it occurs regardless of whether the listener chooses to engage in it or not. Analytic processing of music is more strategic and is exploited at the whim or ability of the listener.

Given that listeners have some flexibility in the application of these various processing strategies, it is not clear how they will be allocated in a melody identification task, which is quite different from the recognition tasks used in previous research. However, if listeners adopt analytic strategies, one would expect to see strong relationships between identification performance and characteristics of individual notes or pairs of notes (i.e., intervals). If listeners tend to adopt holistic strategies, one would expect to see strong positive relationships between identification performance and characteristics that span several notes, and/or provide information about global properties of the melody. The sections that follow classify a range of

musical characteristics as indexes of either analytic or holistic processing. In classifying musical characteristics as analytic or holistic, we extended the philosophy adopted by Peretz and Morais (1987), who considered intervals to be an analytic property and contour to be a holistic property. Therefore, characteristics of individual notes or intervals were considered to be analytic properties and characteristics of longer sequences of notes were considered to be holistic properties.

## Musical Characteristics Associated With Analytic Processing

### INTERVALS

An interval refers to the pitch distance between two consecutive notes. One could argue that intervals should be considered as the fundamental unit of music identification because isolated notes have no inherent meaning in a melody identification task. That is, an isolated tone could conceivably be the first note of *any* melody transposed to the appropriate key; only after the second notes/first interval can some melodies be eliminated from consideration. Intervals can be categorized along many dimensions. Trained listeners can easily distinguish intervals by *size*, and untrained listeners can do so when given a familiar frame of reference (J. D. Smith, Kemler Nelson, & Appleton, 1994). In a corpus of traditional folk tunes, Dowling (1978) found that relatively small intervals were predominant. Therefore, one might expect relatively large intervals to facilitate identification because they are distinctive. Conversely, Narmour (1990) has argued that listeners tend to expect smaller intervals (see also, Schellenberg, Adachi, Purdy, & MacKinnon, 2002). Therefore, small intervals might facilitate identification because they conform to listener expectations.

Intervals also differ in terms of their *direction*: rising to a higher pitch, falling to a lower pitch, or remaining the same. One could predict that falling pitch contours, which often signal the end of a musical phrase, would facilitate identification relative to rising pitch contours. Finally, intervals may also be distinguished by their perceived *consonance*, which is related to how pleasant an interval sounds (see, Krumhansl, 2000, for a discussion). According to Western music theory, perfect intervals (unisons, octaves, fifths, and fourths) are perceived as the most consonant. Major and minor intervals (seconds, thirds, and sixths) are perceived as less consonant, with augmented and diminished intervals being the least consonant. Empirical research generally validates these traditional notions of consonance (Hutchinson & Knopoff, 1978; Kameoka & Kuriyagawa, 1969; Krumhansl, 1990). However, this agreement does not indicate whether the relationship between melody identification and interval consonance will be inverse or direct.

#### DURATION

One might argue that duration is a holistic property, if it is measured *relative* to the lengths of other notes. However, duration will be considered an analytic property because the duration of a single note is independent of the durations of surrounding notes. There are several reasons to suspect that long notes will be locations where identification is especially likely to occur. For example, relatively long notes and/or notes followed by relatively long silent intervals tend to attract attention, thereby creating temporal accents (Handel, 1989; Jones, 1987; Jones & Boltz, 1989). Temporal accents often co-occur with other kinds of musical accents, which would make these locations places where the complex hierarchical structure of the melody was revealed. Put another way, temporal accents may represent locations where musical expectations are either confirmed or violated. This argument implies that relatively long notes provide information that is not available at other serial locations. However, one need not adopt this assumption to argue that listeners will be more likely to identify melodies at relatively long notes. Changes to a melody are easier to detect if the changed note is relatively long (Monahan, Kendall, & Carterette, 1987) or is preceded by a relatively long silent interval (Jones, Boltz & Kidd, 1982). These data suggest that listeners pay more attention to relatively long notes. As such, relatively long should also be relatively familiar, which may help the listener orient to the melody in the way that a relatively familiar intersection helps a traveler orient to an unfamiliar city. According to this argument, long notes do not provide special information; they are simply the locations in the melody that the listener knows best. A third possibility is that relatively long notes provide a “break in the action” during which no new information is given; this might allow the listener to analyze and integrate what has recently been heard.

#### PITCH HEIGHT

Some researchers have hypothesized that extreme notes stand out. Thus, one might predict that notes that were either relatively low or relatively high in pitch would facilitate identification.

### Musical Characteristics Associated With Holistic Processing

#### TONAL FUNCTION

One of the most highly studied aspects of Western music is tonal structure (Krumhansl, 1979, 1990; Krumhansl & Kessler, 1982), which refers to the fact that compositions are typically organized around one central



tone, the tonic (Cuddy, Cohen, & Mewhort, 1981; Krumhansl, 1979). Important tones—those that are closely related to the tonic—are sounded more frequently, for longer durations, and are more likely to mark the ends of phrases than are the other diatonic tones (Jones, 1987). One might argue that tonal function should be considered an analytic characteristic because it is a property associated with a single note. However, this property is determined by the relationship between a single note and the overall musical context. Thus, tonal function is a holistic characteristic because its appreciation depends on the listener's overall interpretation of the melody.

#### CONTOUR PATTERNS

Contour refers to the pattern of rising and falling pitches in a melody; according to some researchers, contour information is more fundamental to music processing than interval information (Dowling, 1978; Dowling & Fujitani, 1971; but see also Dowling & Bartlett, 1981; Edworthy, 1982, 1985). A *reversal* of contour direction refers to a change from a rising pitch line to a falling pitch line, or vice versa. Research has shown that alterations to a previously heard melody are more easily detected when they occur at contour reversals than when they occur at other locations (Dyson & Watkins, 1984). Further, Jones (1987; Jones & Ralston, 1991) has argued that points of contour inflection serve as melodic accents that draw the listeners' attention and have shown that changes in contour direction significantly affect melody perception (Boltz, Marshburn, Jones, & Johnson, 1985; Cuddy et al., 1981; Monahan et al., 1987). Contour reversals are considered an index of holistic processing because their function has been likened to that of corners in visual displays; that is, they establish the broad outlines of the melody (Dyson & Watkins, 1984).

Other contour-based patterns may also facilitate identification by helping to distinguish melodies from one another. However, the music cognition literature does not provide a great deal of guidance with respect to what contour patterns might be associated with melody identification. Some rule-based musical grammars have been proposed (Deutsch & Feroe, 1981), but other researchers have questioned whether these coding schemes are applicable to naturally occurring musical forms (Boltz, 1991). Looking outside the music perception literature, research in serial pattern learning suggests that *runs* and *trills* are more easily encoded than are other patterns (Restle, 1970; Restle & Brown, 1970). A run can be defined as series of consecutive elements (1 2 3 4), whereas a trill consists of an alternating pattern of elements (2 3 2 3). Similar patterns are easy to locate in musical sequences and may facilitate memory and/or identification in this domain, as well.



#### METRICAL ACCENT

The metrical structure of a melody frames the entire piece and has significant effects on melody perception. Palmer and Krumhansl (1987a, 1987b) found that “completeness” ratings for an interrupted melody were significantly higher when they ended on a metrically accented location rather than an unaccented location. Similarly, Dowling, Lung, and Herrbold (1987) found that melodies with interleaved distractor notes were easier to identify if the target notes fell on metrically accented, rather than metrically unaccented, locations. Therefore, metrically accented locations should facilitate melody identification relative to unaccented locations.

#### RHYTHMIC FACTORS

Rhythm—which refers to the specific pattern of durations used to fill a metrical structure—plays an important role in melody perception (Jones, Summerell, & Marshburn, 1987) and influences identification of familiar melodies (Schulkind, 1999). Rhythm is considered to be a holistic property because Jones (1987) has persuasively argued that rhythm guides the listeners’ attention and highlights important aspects of the global structure of a melody. Identification should be more likely when the duration of one note is different from the previous note, especially if the second note is longer than the first. Second, people are best able to reproduce a rhythmic pattern if the intervals between events form simple integer (2:1, 3:1) relationships (Deutsch, 1983; Essens & Povel, 1985; Handel, 1989). Essens (1986) found that reproductions of rhythmic patterns that consisted of subdivisions with noninteger relations (e.g., 2.5:1) were distorted towards an integer ratio (e.g., 2:1). Thus, listeners might be more able to identify melodies when a note is relatively long compared with the previous note and the durations of the notes form an integer ratio.

#### PHRASE BOUNDARIES

Phrase boundaries play an important role in musical structure and therefore may play an equally important role in melody identification. Phrase boundaries serve as the anchors upon which the melody is hung and play a critical role in the generation and confirmation or violation of musical expectancies (Jones, 1987; Narmour, 1990). The beginning of the phrase marks the location where musical tension is set in motion and the end of the phrase is where this tension is often resolved. As well, multiple forms of musical accents (relatively long notes, metrical accents, the tonic and closely related pitches) tend to co-occur at phrase boundaries. Therefore, phrase boundaries may provide information about the global structure of a melody that help the listener either interpret what they have just heard or predict

what will come next. In other words, phrase boundaries may provide information about the global structure of the melody that is difficult to discern within the phrase. This hypothesis suggests that melody identification will be highly correlated with notes both at the beginning and the end of phrases.

Another rationale for predicting that melody identification will be especially likely to take place at phrase boundaries is based on a modified interpretation of the holistic view of melody processing. Rather than cohering into a single monolithic entity, individual notes might cohere into smaller chunks—phrases—that are strung together to form the melody. In fact, many theorists have argued that music is structured hierarchically (Jones, 1987; Lerdahl & Jackendoff, 1983), and research shows that listeners are sensitive to different levels of a musical hierarchy (see Krumhansl, 2000, for a review). It is possible that phrases may be the lowest level of the hierarchy that provides enough information to allow identification. This hypothesis predicts that identification performance will be higher at the end of the phrase than at the beginning because the end of the phrase would mark the completion of an identifiable unit.

### **How Is the Physical Stimulus Reconciled With Its Corresponding Mental Representation?**

Many classes of theories have been proposed to explain how people connect a physical stimulus with its corresponding mental representation. However, few of these theories seem well suited to answer this question for musical stimuli. For example, feature-based theories have been very influential in understanding visual object recognition (Biederman, 1987). However, the features used to identify melodies have yet to be ascertained; in fact, that is one of the major goals of the current work. Prototype theories would also have trouble explaining melody identification (Rosch, 1973). Prototype theory works best for concepts like “birds” or “furniture” that contain many different exemplars linked by family resemblance. Thus, it may help explain why “Rudolph the Red-Nosed Reindeer” is a “better” Christmas Carol than “The First Noel,” but would have difficulty explaining how listeners distinguish these songs from each other.

One reason why these theories may fail to explain melody identification is because they are based on visuospatial stimuli; in contrast, music is an auditory-temporal stimulus. However, researchers have shown extensive interest in studying spoken word identification. Like melodies, spoken words are auditory stimuli, for which the temporal structure is a crucial component of its identity. Therefore, one might expect theories of spoken word identification to provide more insight into melody identification than theo-

ries based on visual stimuli. In particular, the cohort theory of spoken word identification (Marslen-Wilson, 1987; Marslen-Wilson & Tyler, 1980) would seem to be a good approximation to melody identification because no theory places a stronger emphasis on the temporal structure of the to-be-identified stimulus (Jusczyk & Luce, 2002). Briefly, the cohort model relies on bottom-up processing and places special weight on initial sounds of the word (Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, Moss, & van Halen, 1996). Initial sounds activate a cohort of words from the lexicon that share those initial sounds. Identification occurs when enough information is given to distinguish the target from all other entries in the lexicon. The process for melodies might proceed as follows: if a listener hears the first few notes of “Frosty the Snowman,” this song would be a candidate response as would all other songs that shared these initial elements. Each candidate would be retained until information inconsistent with that candidate was presented, and identification would occur when only one candidate remained viable.

In addition to emphasizing the temporal sequence of elements, the cohort model is a good match for melody identification because some lines of evidence suggest that the initial notes of a song are particularly important for melody identification (Schulkind, 2003). Several influential theories of music perception hold that the generation and confirmation or violation of expectancies are central to understanding music (Jones & Boltz, 1989; Narmour, 1990). One would expect the initial notes of the melody to be locations where expectancies were developed. For example, some theorists have argued that the first few notes establish the key of a melody, one of the most important organizing principles of music (Butler, 1992). However, at least one set of data in the literature could be viewed as contradictory to the general claim that the first notes of a melody convey special information. Halpern (1984) showed that similarity ratings of melodies were unrelated to the musical characteristics of the first two notes. However, this result might have obtained because the participants did not actually hear the melodies; rather, they sorted index cards that contained the titles. Therefore, the experiment may not have been particularly musical, an idea reinforced by the fact that genre—a nonmusical characteristic—was the primary determinant of similarity.

The best way to assess the possibility that melody identification makes use of a process similar to that outlined by the cohort model is to examine how identification unfolds across the melody. The cohort model holds that identification will occur when enough information is provided to distinguish a melody from all others stored in memory. This argument implies that identification will be low early in the melody because the first few notes will not provide enough information to distinguish a melody from all other melodies familiar to the participant. Identification should peak at the

point when the target is distinguishable from all others; knowing exactly where this will occur is difficult to predict. However, beyond this point, additional information should prove relatively ineffective. Put more simply, the cohort-based model of melody identification predicts an inverted U-shaped function between identification performance and the number of notes heard. This prediction is simply a restatement of the general principle that initial segments of the melody will be particularly valuable for identification. Examining the relationship between identification and serial position will also be valuable because it will provide normative data about the amount of information listeners need to identify familiar melodies.

## Experiment

The current experiment adopted a novel, two-phase approach to address the question of what factors facilitate melody identification. The first phase consisted of a melody identification task. Unlike past research, which has used either novel melodies (Halpern, 1984) or transformed versions of familiar melodies (Hébert & Peretz, 1997; Schulkind, 1999; White, 1960), the current experiment used unaltered versions of familiar melodies. The target melodies were presented on a note-by-note basis until identification was achieved. The second phase consisted of an in-depth musical analysis of these melodies (see Appendix). The first 12 notes of each melody were coded on a number of musical dimensions (e.g., tonal function, contour, intervals, duration) that were identified as potential mediators of identification performance; Figure 1 displays the first 12 notes and associated musical analysis for “Frosty, the Snowman.” Multiple regression analyses determined how well the coding scheme from the second phase of the experiment predicted the melody identification behavior in the first phase.

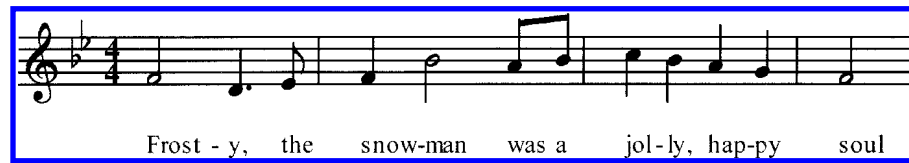
## METHOD

### Participants

Twenty-eight undergraduate volunteers received partial credit toward a course requirement in exchange for their participation. Data on music training was available for only 24 of the 28 participants. These participants averaged 3.50 years ( $SD = 3.69$ ; range 0–13) of formal musical instrument training and 0.50 years ( $SD = 1.06$ ; range 0–4) of performance in organized singing groups; none were music majors.

### Materials

The main criterion for the selection of the target melodies was their familiarity to the participant population. A group of 80 undergraduates was asked to write down a list of songs that they thought “everyone would know.” A second group of 97 undergraduates rated their familiarity with these songs. The 34 songs (see Appendix) used in the experiment



Tonal function										
TF	2	5	6	3	6	2	6	3	2	5
TF-K&K	4.1	5.2	6.4	2.9	6.4	3.5	6.4	2.9	3.7	5.2
Pitch Height										
PH	7	9	14	13	14	16	14	13	11	9
PH-Med	4	2	3	2	3	5	3	2	0	2
Contour										
Direction	1	1	1	-1	1	1	-1	-1	-1	-1
Cnt-Change	3	1	1	3	3	1	3	1	1	1
Cnt-Inflection	1	1	3	3	1	3	1	1	1	3
Interval										
Int-Size	1	2	5	1	1	2	2	1	2	2
Int*Dir	1	2	5	-1	1	2	-2	-1	-2	-2
PMMA	2	3	4	2	2	3	3	2	3	3
Int-K&K	285	275	245	285	285	275	275	285	275	275
Int-H&K	.49	.24	.03	.46	.46	.17	.16	.46	.20	.22
Duration										
#Beats	.5	1	2	.5	.5	1	1	1	1	2
Rhythmic factors										
Rel-Dur	.33	2	2	.25	1	2	1	1	1	2
Interval ratios	2	2	2	2	2	2	2	2	2	2
Successive ratios	1	1	1	1	1	1	1	1	1	1
Meter										
Meter-L&J	1	3	2	2	1	3	2	3	2	3
Meter	1	4	2	2	1	4	2	3	2	4
Phrasing										
Phrase placement	3	2	1	1	2	3	4	3	2	1
Contour patterns										
Alternations	0	0	0	0	1	0	0	0	0	0
Runs	0	0	1	0	0	0	0	0	1	2
Pairs	0	0	0	0	0	0	0	0	0	0
Serial Position										
SP	3	4	5	6	7	8	9	10	11	12
SP <sup>2</sup>	9	16	25	36	49	64	81	100	121	144
SPu	1	2	3	4	5	5	4	3	2	1

Fig. 1. The first 12 notes of “Frosty, the Snowman” and its associated musical analysis. The analysis is presented for notes 3–12, as these were the notes used in subsequent analyses.

were selected because a majority of the second group claimed that they “easily knew most of the words.” The stimuli incorporated a wide range of keys (e.g., C, G, F, Eb, Ab), both simple (2/4, 3/4, 4/4) and complex meters (6/8), and were drawn from several genres (e.g., patriotic songs, Christmas Carols, children’s songs, folk music, “pop” music, movie songs, Broadway show tunes).

A published sheet music arrangement was obtained for each song. The first 20 notes of the melody line of the published arrangements were transferred to a Macintosh IIsx computer using the Songworks (Ars Nova) software package. This program was also used to

present the melodies, which were heard over headphones with a piano timbre. The first two authors adjusted the tempo of each melody to sound “natural” ( $M = 117.62$  quarter notes per minute,  $SD = 26.83$ ; range 75–150). Presentation order was randomized and played in either a backward or forward order. Order of presentation did not influence initial data analyses, and thus was dropped from the analyses reported in the results section.

### Musical Analysis

The first two authors conducted the musical analysis. Each coder analyzed the first 12 notes of each song by using the dimensions described below (see Figure 1). Although identification data were available for up to 20 notes, only the 3rd–12th notes were analyzed. First notes were not analyzed because identification data were not collected for the first note in isolation; second notes were not analyzed because the second note failed to yield data values for the contour change and local pattern measures (see below). The analysis ceased at note 12 because the vast majority of melody identifications (87%) occurred by this point. Discrepancies between the two coders were resolved by discussion. The following predictors were entered into the regression model. When appropriate, different measures of the same general construct are distinguished by abbreviations in parentheses.

### Pitch Factors

#### Tonal Function

Tonal function (*TF*) was coded in terms of perceived centrality to the key based on traditional Western music theory. In decreasing order of centrality, this system distinguished between tonic, dominant, mediant, leading tone (because its presence strongly implies the tonic), other diatonic tones, and nondiatonic tones. This coding scheme implied a linear relationship between different types of tones, which may not reflect the actual relationship (Krumhansl, 1979, 1990). Therefore, goodness of fit ratings for the various tones in the chromatic scale—as reported by Krumhansl and Kessler (1982)—were also used (*TF-K&K*).

#### Pitch Height

Pitch height (*PH*) was coded using either the lowest tone ( $A_{b3}$ ) or the median (*PH-Med*) tone in the corpus ( $G_4$ ) as an anchor. The anchor notes for these measures were assigned a value of 0, with all other notes coded in terms of the number of semitones from the anchor. All values for *PH-Med* were positive regardless of whether the note was higher or lower in pitch than the anchor.

#### Contour

Several aspects of contour were coded. First, we considered whether a note created a rising (+1), falling (-1), or neutral (0) interval (*Dir*). Next, we coded whether a note created a change in pitch contour direction (*Cnt-Chng*). For example, in the sequence  $C_4 E_4 D_4 C_4$ , the  $D_4$  changes what had been a rising pitch line to a falling pitch line. Notes were coded as either continuing the pitch trajectory (1), or altering the pitch trajectory (3); an intermediate value (2) was given to unisons because the effect of unisons on pitch trajectory was indeterminate. Finally, points of contour inflection (*Cnt-Inf*) were also coded. This factor differed from the other contour measures in that it coded information about the relationship of the intervals preceding and following the coded note. If the intervals preceding and following a given note differed in direction—as with the  $E_4$  in the example just presented—the note was considered a point of contour inflection and was coded with a 3. If the preceding and following intervals were in the same direction (e.g.,  $D_4$  above), the note was not considered a point of contour inflection and was coded as a 1. Owing to their ambiguity, notes followed by unisons were coded with an intermediate value of 2. If a note was preceded by a unison, the following interval was compared with the direction of the closest, preceding, nonunison interval. For example the second  $E_4$  in the sequence  $D_4 E_4 E_4 C_4$  was coded as a point of contour inflection because, whereas the  $D_4-E_4$  interval is rising, the  $E_4-C_4$  interval is falling.

### *Interval*

Interval size was coded in semitones. For one measure, rising and falling intervals were treated identically (*Int-Size*), whereas another incorporated information about direction as well (*Int\*Dir*). The perceived consonance of different intervals was also considered as a factor. Again, one method was based on traditional Western music theory. In order of decreasing consonance, intervals were identified as perfect, major, minor, and augmented (*PMMMA*). Published judgments of interval consonance were also taken from both Kameoka and Kuriyagawa (1969; *Int-K&K*) and Hutchinson and Knopoff (1978; *Int-H&K*).

### *Temporal Factors*

#### *Duration*

Notes were coded in terms of their length relative to both the beat period (*#Beats*) and the preceding note (*Rel-Dur*). For example, if a song was written in 4-4 meter, a quarter note preceded by an eighth note would be coded as a 1 for *#Beats*, and a 2 for *Rel-Dur*. Similarly, notes were also coded in terms of whether their duration formed an integer relationship between either the beat period (*interval ratios*) or the preceding note (*successive ratios*).

#### *Meter*

One system for coding metrical structure was based on Lerdahl and Jackendoff's (1983) model (*Meter-L&J*). Three levels were coded: strong (first/third beat of the measure), weak (second/fourth beats of the measure), and weaker (half-beats). Notes that spanned more than one beat were coded using the beat on which the note began. Songs that did not contain a third or fourth beat were coded in the same way. For songs written in 2/4 meter, any note beginning on the first beat was coded as strong, any note beginning on the second beat was coded as weak, and any note beginning on a half-beat was coded as weaker. For songs written in 3/4 meter, any note beginning on the first or third beat was coded as strong, any note that began on the second beat was coded as weak, and all other notes were coded as weaker. For songs written in 6/8 meter, notes that began the measure were coded as strong; notes that began on the second beat (the fourth eighth-note) were coded as weak, and all other notes were coded as weaker.

A second system (*Meter*) separated the first and third beats into separate categories (e.g., strongest and strong), followed by the second/fourth beats (weak) and half-beats (weaker). For songs written in 2/4 meter, any note beginning on the first beat was coded as strongest, any note beginning on the second beat was coded as weak, and any note beginning on a half-beat was coded as weaker. For songs written in 3/4 meter, any note beginning on the first beat was coded as strongest, any note beginning on the third beat was coded as strong, any note that began on the second beat was coded as weak and all other notes were coded as weaker. For songs written in 6/8 meter, notes that began the measure were coded as strongest; notes that began on the second beat (the fourth eighth-note) were coded as weak, and all other notes were coded as weaker. The second system was adopted to evaluate the possibility that the first beat of the measure serves as the anchor around which the rest of the measure is organized.

### *Melodic Factors*

#### *Phrase Placement*

Because the determination of phrase boundaries is difficult to assess objectively, initial assessments were made using verbal phrase boundaries from the song lyrics. Notes were then coded as either forming the beginning or ends of a phrase (1), or in terms of the number of notes separating the note from a phrase boundary (e.g., the second and penultimate notes of a phrase were coded as 2, etc.). The placement of phrase boundaries was corroborated by examining whether notes at phrase boundaries tended to be marked by multiple pitch (tonic triad members, contour changes, etc.) and temporal (relatively long notes, met-



rically accented notes) accents. A one-way analysis of variance (ANOVA) showed that accent density differed across placement in the phrase,  $F(3, 230) = 8.05$ ,  $MSE = 728.8$ ,  $p < .01$ . Tukey tests ( $\alpha = .05$ ) indicated that notes at the beginnings and ends of phrases created pitch and temporal accents more frequently than notes in the middle of phrases, but that notes in the middle did not differ from one another.

#### Local Patterns

Three types of local patterns were considered. Two (alternations and runs) were selected on the basis of previous work (Restle, 1970; Restle & Brown, 1970). The third (pairs) was selected because it was a recurring pattern in the selected corpus. *Alternations* were defined as 2 consecutive changes in pitch contour direction. As such, alternations were similar to trills but did not require that the exact same notes repeat. For example, in the sequence  $C_4 E_4 C_4 D_4$ , the  $D_4$  would be considered the completion of an alternation and would be coded as a 1. If the next note created another contour change (e.g.,  $C_4$ ), that note ( $C_4$ ) was coded as a 2, and so on. In this way, the alternation variable coded both the presence and length of an alternating sequence. A *run* was defined as four consecutive notes that shared either a rising or falling pitch trajectory; unisons were not considered part of a consistent trajectory. The fourth note in the sequence was coded as a 1; if the next note continued the same trajectory, it was coded as a 2, and so on. Finally, a *pair* was defined as any repeated combination of two notes sharing an easily recognizable pitch relationship. For example, the first six notes of "Twinkle-Twinkle" ( $F_4 F_4 C_5 C_5 D_5 D_5$ ) contained three pairs of unisons. The first, second, third, and fifth notes were coded with a zero; the fourth and sixth notes were coded with a 1 to indicate that they completed the repetition of a pair).

#### Serial Position

Serial position (*SP*) was recorded and entered into the analysis to examine whether early portions of the melody were differentially important for identification. Three different measures of *SP* were recorded to examine whether the relationship between *SP* and identification was linear, exponential, or U-shaped. *SP* was the serial position of the note in the melody, and  $SP^2$  was the square of this value. *SPu* was created by assigning a value of 1 to *SPs* 3 and 12, a value of 2 to *SPs* 4 and 11, a value of 3 to *SPs* 5 and 10, and so on.

#### Procedure

Participants were tested individually. They were told that they would be asked to identify a series of familiar songs in as few notes as possible. For each song, listeners initially heard the first two notes of the song. They were instructed to give one of the following five responses: (1) "I have no idea"; (2) "It sounds familiar but I am not certain enough to make a guess"; (3) "I am not sure but it might be \_\_\_\_\_"; (4) "I am pretty sure that it is \_\_\_\_\_"; or (5) "I am sure that it is \_\_\_\_\_". The five response options were printed on a sheet of paper placed on the desk in front of the participant.

If they failed to identify the song with certainty (i.e., response category "5") after the first two notes, they heard the first three notes of the song. Notes were added one by one until either the tune was named with certainty, or the 20th note was reached. If the participant failed to identify the song by the 20th note, s/he was instructed that a new song was to be started, but was not told the title of the unidentified tune. Feedback was only provided if the "certain" response (i.e., option "5") was used. If a "certain" response was correct, the participant was told, "That is correct. Now we will begin the next song." If a "certain" response was incorrect, the participant was told, "I'm sorry, that response is incorrect," and the next trial for that song was played. Participants were required to respond within 10 seconds of the end of a given trial; the experimenter monitored the 10-second window with a stopwatch. One practice trial ("Old McDonald Had a Farm") was given to familiarize participants with the task. At the end of the experiment, participants were given a list containing all 34 songs used in the experiment. They were asked to indicate if any of the target songs were unfamiliar; these data points (< 2%) were removed from all analyses.

## RESULTS

Multiple regression analyses were conducted to determine which predictor variables were associated with identification performance. Identification performance was operationalized in terms of the conditional probability of identifying the song at each note. This measure was defined as the number of participants who identified a song on a given note divided by the number of participants who had yet to identify the song. For example, if 7 listeners identified a song after hearing the first 2 notes, and an additional 7 identified the song after hearing the first 3 notes, the conditional percent correct for note 2 would be  $7/28 = .25$ , whereas the conditional percent correct for note 3 would be  $7/(28-7) = 7/21 = .33$ . This dependent measure was chosen in lieu of raw percent correct because, whereas raw percent correct would by definition decrease as more and more notes were heard (and more and more participants had successfully identified the tune), the conditional percent correct was free to vary from 0% to 100% on every trial.<sup>1</sup>

**Initial Regression Analysis**

The 12 notes from 34 songs yielded a total of 408 data points. However, data were only used for SPs 3–12. As described earlier, data for SP 1 were not used because identification data were not collected for the first note in isolation and data for SP 2 were eliminated because they failed to yield data values for the contour change and local pattern measures. The loss of these 68 data points is tolerable, given that no participants were able to identify any of the songs before SP 3. In addition, data points were eliminated if 5 or fewer participants had yet to identify the song. This procedure, which resulted in the loss of 120 data points, was adopted because using data when only a handful of listeners had yet to identify a song tended to inflate the conditional percent correct. Regression analyses were conducted using the “Stepwise” regression technique in SPSS (Release 10.0.5). For each step, the predictor with the lowest significant  $p$  value was entered into the model ( $\alpha = .05$ ). Upon entry, all previously entered factors were re-examined to determine whether or not they should be retained. The model was considered complete when no additional predictors entered the model.

The results of the regression analysis are summarized in Table 1. Four variables entered into the overall regression equation: SPU, phrase placement, alternations, and #Beats,  $F(4, 228) = 24.44$ ,  $MSE = 291.01$ ,  $p < .01$ .

1. A regression analysis conducted with raw percent correct as the dependent measure. Two significant predictors entered into the regression equation: SPU and meter,  $F(2, 230) = 15.50$ ,  $MSE = 123.32$ ,  $p < .001$ . Overall  $R^2$  for the model was .119, and the two factors uniquely explained 10.8% of the variance in identification performance.

TABLE 1  
Summary of Output for Regression Analyses

Model	Predictor Variables						Variance Explained	
	SPu	Phrase	Altern	#Beats	Meter	Pairs	R <sup>2</sup>	Unique R <sup>2</sup>
Overall	.185	.059	.037	.019			.300	.214
Worse	.085	.025	.025				.135	.119
Better	.091	.038			.023	.030	.182	.169
Random1	.224	.055	.034		.026		.339	.239
Random2	.265	.082	.064				.410	.351
Random3	.205	.069			.039		.313	.249
Random4	.233	.099					.332	.239

NOTE—Values in the table indicate the incremental change in  $R^2$  that accompanied the entrance of each factor into the model. The larger this value, the earlier the factor entered the model.

The beta coefficient for the SPu variable indicated that identification performance was higher in the middle of the presented melodies (SPs 5, 6, and 7) than at the ends (SPs 3–4, 10–12). In other words, identification performance traced an inverted U-shaped function across SP. The beta coefficients for the phrase placement, alternations, and #Beats factors indicated that better identification performance was associated with notes at or near phrase boundaries, notes that completed alternations, and relatively long notes, respectively.

### Individual Differences

A number of subsidiary analyses were conducted to examine the reliability of the results in the main analysis and to address some of its limitations. One possible limitation of the regression analysis was that the skill of the participant population may have decreased across SPs because the most proficient participants dropped out of the analyses earlier than the less proficient participants because they identified the songs sooner. This issue was examined by conducting separate regression analyses for the more and less proficient participants. A median split of the data produced a group of *better* ( $n = 15$ ) and *worse* ( $n = 13$ ) identifiers; the two groups were not equal in size because of ties. The better group ( $M = 6.83$ ,  $SD = 0.61$ ) identified songs in fewer notes than the worse group ( $M = 8.94$ ,  $SD = 1.04$ ). This difference was significant,  $t(26) = 6.66$ ,  $SEM = .32$ ,  $p < .01$ , but may actually underestimate the difference between the groups because the data included only songs that were identified correctly.

The procedure for these regression analyses was identical to the used for the main analysis with one exception; because there were fewer participants in these analyses, data points were eliminated only if three or fewer

participants had yet to identify the song. Three variables entered into the overall regression equation for the *worse* identifiers (see Table 1): SPu, phrase placement, and alternations,  $F(3, 215) = 11.01$ ,  $MSE = 0.02$ ,  $p < .01$ . This pattern is very similar to that produced by the main analysis, the only difference being the exclusion of the #Beats variable in the analysis. Four variables entered into the overall regression equation for the *better* identifiers (see Table 1): SPu, phrase placement, pairs, and meter,  $F(4, 148) = 24.44$ ,  $MSE = 291.01$ ,  $p < .001$ . This model included the first two predictor variables from the main analysis, but differed from the main analysis in its exclusion of #Beats and alternations, and its inclusion of the pairs and meter variables.

Individual differences in the target melodies were also considered. Although it might have been desirable to pursue this question by conducting separate regression analyses for each song, the number of data points produced by each song (i.e., the 10 SPs) was insufficient given the number of predictor variables included in the main analysis. Therefore, this issue was addressed by randomly selecting half of the data points for analysis; this process was repeated four times. As can be seen in Table 1, the four models based on randomly selected cases were largely in agreement with each other and with the regression equation based on the data set as a whole. The first and second factors to enter each model were SPu and phrase placement, respectively. For two of the models, alternations entered as a third factor. Two of the models also included meter as a factor: one as the third factor in the model, the other as the fourth factor.

### Participant Analyses

As can be seen in Table 1, the regression analyses converged on a relatively restricted set of predictor variables. Every model included the SPu and phrase placement variables as the first and second predictors to enter the model. Four of the models included alternations as the third factor and three models included meter as a predictor; #Beats and pairs were included in one model each. However, the regression analyses that produced these models can be criticized because the units of analyses (notes from each song) were not independent of one another. In contrast, the performance of each participant was independent of all other participants. Therefore, the data were re-analyzed with participants as the unit of analysis to corroborate the results of the note-based regression analyses. Rather than examine all of the variables included in the regression models, these analyses focused on those factors that were found to be significant predictors of identification.

For each participant, percent correct was calculated separately for each level of the five predictor variables implicated by the regression analyses.

For example, the #Beats variable was broken down into three categories: tones that were less than one beat in length (e.g., SPs 3, 6, and 7 in “Frosty, the Snowman”), tones that were equal to one beat in length (e.g., SPs 4, and 8-11 in “Frosty, the Snowman”), and tones that were greater than one beat in length (e.g., SPs 5 and 12 in “Frosty, the Snowman”). The number of notes that fell into each category was counted; these numbers varied from participant to participant because only those notes actually heard by a participant were included in the analysis; that is, if a participant identified “Frosty, the Snowman” at SP 7, SP 8–SP 12 were dropped for that participant. The number of notes within each category that yielded identification was also tabulated. The proportion of notes in each category that yielded identification was calculated by dividing the number of notes that yielded identification by the total number of notes in the category. For example, across the entire experiment, one participant heard 96, 53, and 13 notes that were less than, equal to, and greater than one beat, respectively. The same participant identified 13 songs after hearing a note that was less than one beat, 10 songs after hearing a note that was equal to one beat and 5 songs after hearing a note that was greater than one beat. Thus, the percentage of notes that yielded identification was 14% (13/96) for notes that were less than one beat, 19% for notes that were equal to one beat, and 28% for notes that were greater than one beat.

For all analyses reported in this section, the degrees of freedom were adjusted by using the Greenhouse-Geisser procedure when appropriate; all reported degrees of freedom represent the unadjusted values. The  $p$  value for all pairwise comparisons was adjusted using the Bonferroni correction.

A repeated-measures ANOVA was conducted to examine differences between levels of the phrase placement variable. The data were collapsed into four categories: phrase boundaries (either the beginning or the end) and notes occupying positions 1, 2, and “3 or more” notes from a phrase boundary (see Figure 2). Notes at the beginnings and ends of phrases were collapsed into a single category because both are places that tend to be highlighted by pitch and temporal accents. Analyses indicated that there were no differences in identification performance at the beginnings of phrases relative to the ends of phrases. The overall ANOVA was significant,  $F(3, 81) = 33.43$ ,  $MSE = 0.05$ ,  $p < .01$ . Post-hoc comparisons revealed significant differences between all levels of the independent variable with the exception of the comparison between 2 and “3 or more” notes from a phrase boundary. This suggests that notes at phrase boundaries and, to a lesser degree, notes close to phrase boundaries facilitated identification relative to notes in the middles of phrases.

The alternations variable was collapsed into three categories. Notes that did not complete an alternation, notes that completed a four-note alternation (four notes were required to complete the pattern so this represented

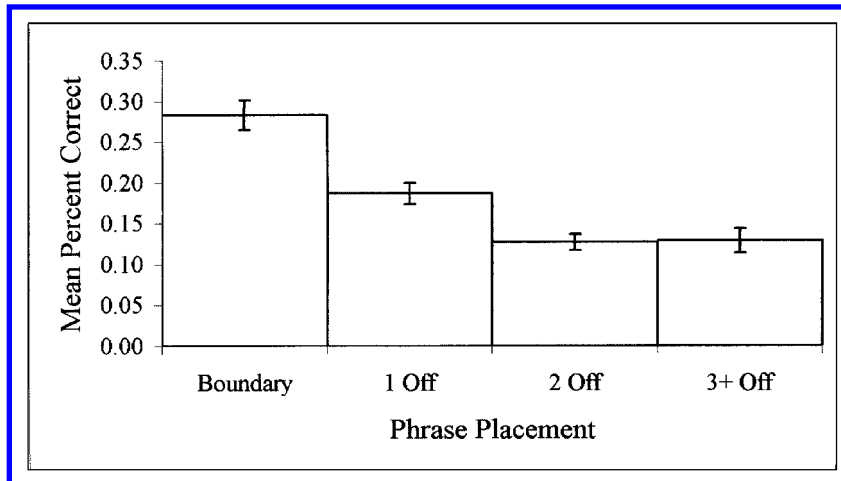


Fig. 2. Mean percent correct for notes located at phrase boundaries, one note removed from a boundary, two notes removed from a boundary, and three or more notes removed from a boundary.

the shortest alternation sequences), and notes that completed longer alternation sequences of between 5 and 7 notes (see Figure 3). The repeated-measures ANOVA was significant,  $F(2, 54) = 7.64$ ,  $MSE = 0.04$ ,  $p < .01$ . Post-hoc comparisons revealed that performance in the “no alternation” condition was significantly worse than in either of the alternation conditions but that the two alternation conditions did not differ from one another. This result indicates that alternations are a valuable feature for identification, but that extended alternations do not confer any particular advantage over shorter alternations.

The #Beats variable was recoded into three levels: less than one beat, equal to one beat, and greater than one beat (Figure 4). The repeated-measures ANOVA was significant,  $F(2, 54) = 32.46$ ,  $MSE = 0.06$ ,  $p < .01$ . Post-hoc comparisons revealed significant differences between all levels of the independent variable. Thus, identification tended to occur when the participants heard relatively long notes.

A repeated-measures ANOVA was conducted to examine differences between the levels of the meter variable (Figure 5). The overall ANOVA was significant,  $F(3, 81) = 22.16$ ,  $MSE = 0.04$ ,  $p < .01$ . Post-hoc comparisons revealed the following patterns. The first beat of the measure yielded significantly better performance than any other location. The third beat of the measure was not significantly different from either the second/fourth beat or from the half-beat condition, but the latter two were different from each other. This pattern was partially consistent with expectations in that the first beat of the measure was by far the best location for melody identifica-

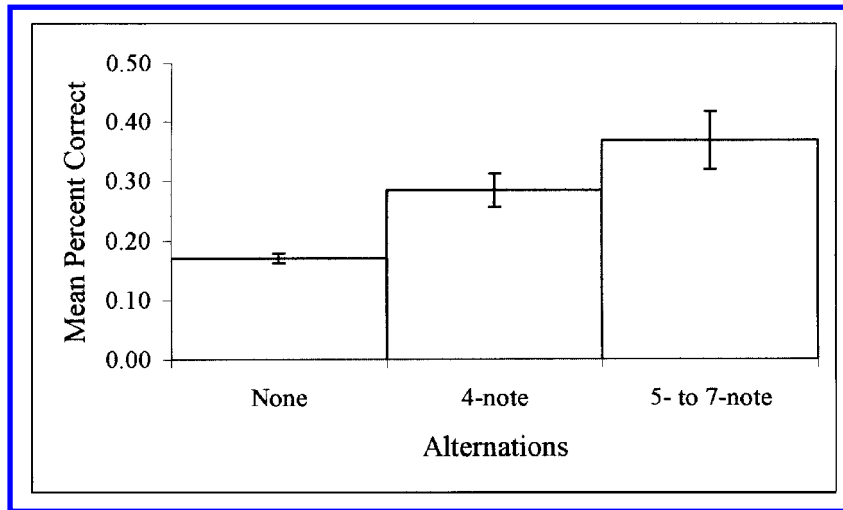


Fig. 3. Mean percent correct as a function of alternations; none refers to a note that was not part of an alternation, 4-note refers to a note that completed a four-note alternation, 5- to 7-note refers to a note that completed an alternation of five or more notes.

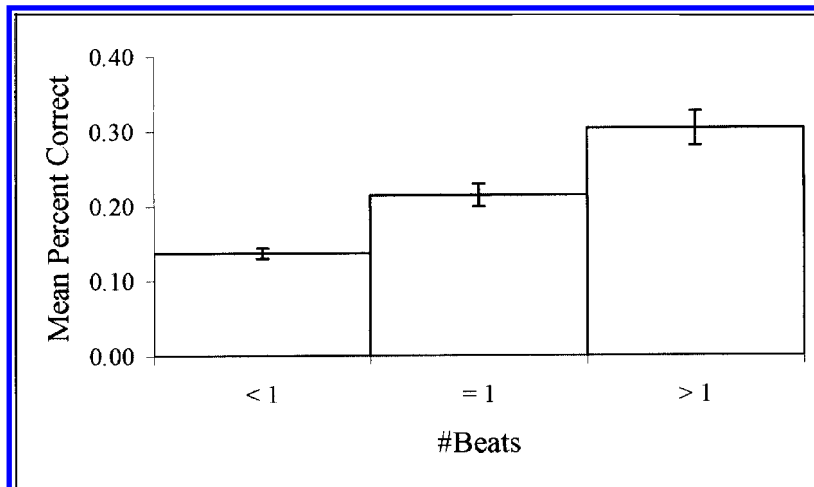


Fig. 4. Mean percent correct as a function of duration in beats; duration was categorized as either less than, equal to, or greater than one beat.

tion. However, the relatively poor performance for the third beat of the measure was surprising, particularly in light of Lerdahl and Jackendoff's (1983) work.



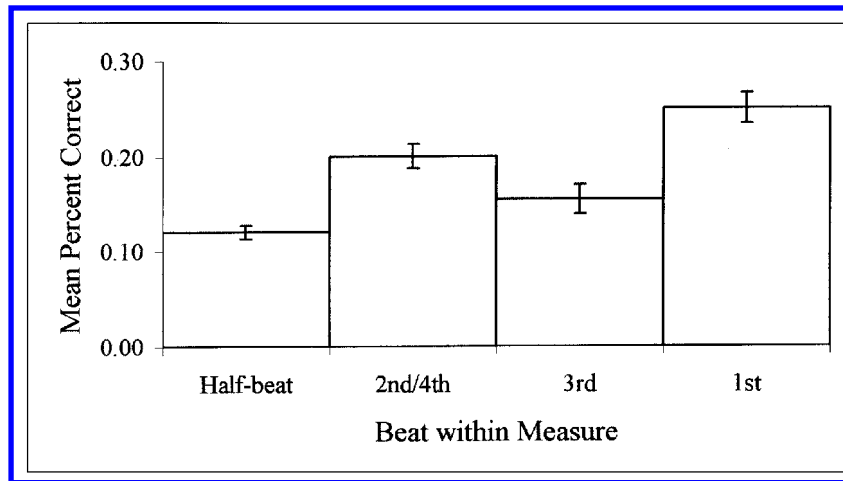


Fig. 5. Mean percent correct as a function of meter; meter was categorized as falling on either a strongest beat (1st beat in the measure), a strong beat (3rd beat in the measure), a weak beat (2nd or 4th beat in the measure), or a half-beat.

Mean identification was also calculated for notes that were not part of a pair ( $M = 0.18$ ;  $SD = 0.01$ ) and those that were part of the first iteration of a pair ( $M = 0.29$ ;  $SD = 0.05$ ). A repeated-measures ANOVA revealed a significant difference between the two levels of this variable,  $F(1, 27) = 79.25$ ,  $MSE = 0.04$ ,  $p < .01$ .

Finally, a repeated-measures ANOVA was conducted to examine the relationship between SP and identification performance. The expected, inverted U-shaped pattern was observed (Figure 6). The overall ANOVA revealed a significant effect of SP,  $F(9, 225) = 9.23$ ,  $MSE = 0.02$ ,  $p < .01$ . Trend analyses indicated the presence of significant quadratic,  $F(1, 25) = 31.42$ ,  $MSE = 0.02$ ,  $p < .01$ , cubic,  $F(1, 25) = 23.68$ ,  $MSE = 0.02$ ,  $p < .01$ , and fifth order,  $F(1, 25) = 5.41$ ,  $MSE = 0.02$ ,  $p < .03$ , trends. These results indicate that identification was most often achieved after moderate amounts of information.

### Confidence Ratings

On average, the songs were identified by the sixth note ( $M = 6.35$ ;  $SD = 0.66$ ). However, the number of notes required to identify familiar melodies may have been overestimated with this method because trials were not terminated until the participant responded with certainty. The participants may have identified the song, albeit with some hesitation, significantly sooner. Therefore, we analyzed the confidence ratings to see how identification unfolded across serial position. For 51.6% of experimental trials, the transition be-

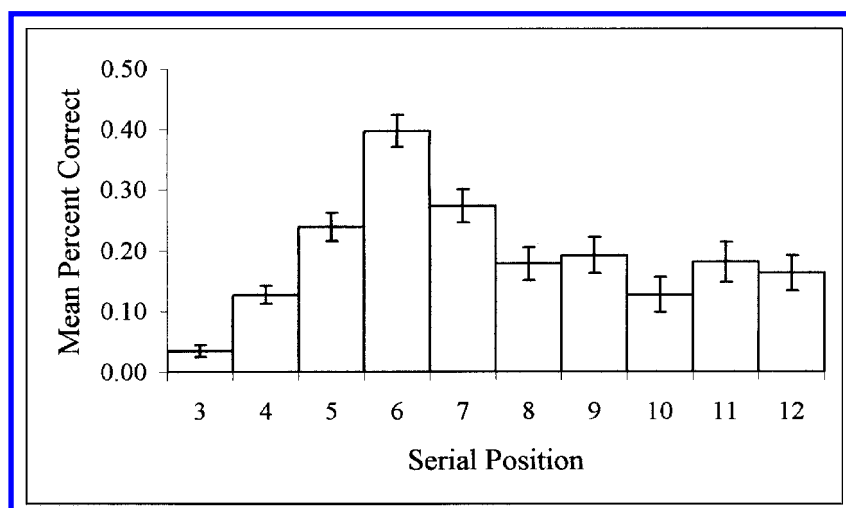


Fig. 6. Mean percent of songs identified as a function of serial position; the number of songs identified at each serial position was divided by the total number of notes heard at each serial position.

tween the inability to generate a guess (a confidence rating of 1 or 2) and complete certainty (a confidence rating of 5) occurred within a single note. For 87.0% of trials, the transition from uncertainty to complete certainty occurred within two notes. These data contradict the hypothesis that the method yielded inflated estimates of the notes required to identify a melody and suggest that the process of identifying a melody is akin to the proverbial light bulb suddenly being lit; that is, melody identification is a relatively quick, “all-or-none” behavior. Almost 90% of the time (Table 2), the participants were either unable to make a guess (1 or 2) or were completely certain of their response (5).

TABLE 2  
Percentage of Responses in Each Confidence Rating Category  
for Correct and Incorrect Responses

Confidence Rating	All Responses		Responses With Titles	
	Correct	Incorrect	Correct	Incorrect
No idea	49.31	—	—	—
Not sure enough to guess	24.38	—	—	—
Might be	3.68	1.21	13.98	4.62
Pretty Sure	5.54	0.69	21.07	2.61
Certain	14.70	0.49	55.86	1.87

NOTE—“Responses with Titles” are those responses for which the participant provided the title of a specific song.

In addition, if the data are restricted to those instances when the participants produced a specific song title, more than half of the participants' responses were given with complete certainty, and very few of these responses were errors. These data are reminiscent of those observed in the insight problem-solving literature (Metcalf & Wiebe, 1987). However, a stronger evaluation of the relationship between melody identification and insight problem solving would require an experimental paradigm that included continuous (rather than note-by-note) melody presentation and online measures of confidence so that more precise information about the timing of changes in confidence ratings could be collected.

## General Discussion

The current experiment was designed to answer two questions. What are the musical features that listeners use to identify familiar melodies? And, how is the physical stimulus compared to its corresponding mental representation? The implications of the data for each of these major questions are discussed in turn.

### WHAT ARE THE FUNDAMENTAL UNITS FOR MELODY IDENTIFICATION?

Five musical characteristics were significant predictors in at least one of the regression analyses. Of these, the placement of a note within a musical phrase was the most consistent predictor and also explained the most variance (Table 1). As predicted, the data indicated that identification performance was highest at phrase boundaries. Because no differences were found between the beginnings and the ends of phrases, the data were inconsistent with the claim that phrases represent the smallest identifiable chunk in the musical hierarchy. However, the data are consistent with the argument that phrase boundaries facilitate identification because these locations reveal information about the global structure of the melody (Jones, 1987; Jones & Boltz, 1989). This hypothesis is also supported by the previously discussed analysis showing that phrase boundaries tended to encompass significantly more pitch and temporal accents than did notes in the middle of phrases. Thus, it appears that tune identification requires a global understanding of melodic structure that is most easily perceived at phrase boundaries.

Two contour-based characteristics were also associated with high levels of identification performance. Notes that completed consecutive alternations between rising and falling pitch contours were a significant predictor in half of the regression models; the pairs variable made a small but significant contribution to one of the regression models. These results may have

obtained because these musical ornaments are easily encoded (Restle, 1970). Another possibility is that these musical ornaments are easily remembered because they are salient or relatively atypical contour patterns (e.g., D. A. Smith & Graesser, 1981) that stand out from the surrounding musical context. This hypothesis is supported by the fact that the experimental corpus included a relatively small proportion of notes that completed alternation (12.1%) and pair patterns (2.4%).

The last two musical variables that entered the best-fit models were both temporal in nature. The presence of the duration relative to the beat and meter variables suggests that identification was most likely to occur at long notes and metrically accented locations. One explanation for the significance of the duration factor is that long notes provide listeners with a temporal pause that allows them to analyze recently heard musical elements. This hypothesis is tenuous because the method used in the experiment incorporated a 10-s response interval after each exposure to the melody; participants could have easily used this time to reflect on the music they had just heard regardless of the length of the last note. A second possibility is that temporally accented notes are locations where other forms of accenting are likely to occur. There are two reasons to question this hypothesis. First, this argument implies a correspondence between pitch and temporal accents, which should have resulted in at least some of the pitch accents attaining statistical significance; none were significant predictors in any model. Second, this argument is redundant with one made for the phrase boundary variable. Given that the phrase boundary variable was a much more consistent and robust predictor, one would have to conclude that the significance of the temporal accents indicates that they contribute something above and beyond coupled accents. This added value may have to do with the allocation of attention. According to Jones (1987) listeners direct more attention to relatively long notes. Therefore, as argued in the introduction, relatively long notes may be processed more at perception and may be more familiar at identification, which helps the listener orient to the overall musical structure.

More generally, the data are consistent with the idea that melody identification tasks recruit holistic rather than analytic processing. Of the five musical characteristics that were significant predictors in at least one regression model, four (phrase placement, alternations, pairs, and meter) would be classified as holistic in nature. That is, they coded information about the overall temporal and pitch shape of the melody rather than characteristics of individual notes or intervals. The alternations and pairs factors were both contour-based characteristics, which previous research has associated with holistic processing (Peretz & Morais, 1987; Peretz et al., 1987). The phrase boundary data suggest that musical phrases may be processed as unified, coherent entities or *gestalts*, a phenomenon that has been

observed in many recognition memory tasks (Deutsch, 1980; Dowling, 1973). The meter of a melody should be considered a holistic property because it provides the global temporal framework of a piece. The only analytic property that made a significant contribution to identification performance was duration. The significance of this factor is consistent with expectancy-based models of music perception (Jones, 1987; Narmour, 1990). According to these models, one might expect duration to be positively correlated with phrase boundaries and metrically accented locations. Therefore, this result is not surprising given the previously discussed findings. Further, the significance of this factor indicates that holistic processing does not occur to the exclusion of analytic processing. Rather, it would seem that holistic processing is dominant for familiar melody identification.

Although many of the data were consistent with previous work, several unexpected results merit further consideration. For example, whereas tonal structure plays a central role in the perception of Western music (Dowling, 1978), neither tonal function variable was a significant predictor in any model. This result may have obtained because the song corpus included very few nondiatonic tones ( $\approx 1.5\%$ ), which would have been the easiest for the untrained participants to distinguish. To the extent that this corpus is representative of Western music in the culture, one would conclude that tonal function does not contribute much to melody identification outside of the laboratory. However, different results might be obtained if more complex musical styles (i.e., classical music), or more sophisticated participants were sampled.

Another surprising result was that temporal factors (metrical accent and duration relative to the beat) contributed more to melody identification than did pitch factors (no significant predictors). Previous research seemed to show that manipulations of the temporal structure of melodies had relatively little effect on identification compared with manipulations of pitch structure (Hébert & Peretz, 1997; White, 1960; for a different view, see Schulkind, 1999). These data are difficult to interpret because the temporal and pitch manipulations in these experiments were not equivalent. Whereas the pitch manipulations (setting each note to the same pitch) completely eliminated the pitch pattern, the temporal manipulations (setting each note to the same duration or applying the rhythmic pattern of one song to another pitch pattern) left long stretches of the rhythmic pattern unaltered (Expt. 1:  $M = 4.6$ ,  $SD = 2.8$ ; Expt. 2:  $M = 2.8$ ,  $SD = 1.1$ ). Thus, these experiments overestimated the importance of pitch information relative to temporal information, a conclusion that was borne out by the current data. However, it should be noted that even the temporal factors contributed relatively little to melody identification when compared with factors that were not strictly tied to either pitch or rhythm.

Finally, the relative importance of contour information at the expense of interval information is surprising given that past research has shown that

listeners are more likely to rely on interval information with familiar melodies and long retention intervals (Dowling & Bartlett, 1981; Edworthy, 1982, 1985). This discrepancy is probably related to the use of different operational definitions of familiarity and retention interval. In previous work, “familiar” melodies were novel before the experiment, but had been repeated between 5 and 15 times during the experimental session. As well, the retention interval was on the order of minutes. The stimuli in the current experiment had been encountered many more times and over much longer retention intervals. As well, previous work employed a recognition paradigm rather than an identification paradigm.

In sum, the fundamental units for melody identification appear to be contour-based, holistic properties that stretch across several notes up to and including phrases. The reliance on holistic rather than analytic processing has important implications for the way musical knowledge is represented in memory. Specifically, these data suggest that melody identification is a procedural memory skill because musical knowledge does not appear to be *compositional* (N. J. Cohen & Eichenbaum, 1993); that is, melodies are not built up from their constituent components (individual notes). One might also look to the previously discussed “all-or-none” nature of melody identification as support for this claim. Given that different brain systems are believed to process procedural and declarative information (N. J. Cohen & Eichenbaum, 1993), these data might influence how neuroscientists search for the neural substrates associated with musical behavior (Peretz & Kolinsky, 1994; Peretz et al., 1994).

The holistic processing bias is also important because it can explain one puzzling aspect of the current data. Several musical features that are easily perceived by relatively untrained listeners—for example, tonal function (Dowling, 1978) and melodic contour (Dyson & Watkins, 1984; Jones & Ralston, 1991)—failed to emerge as significant predictors of melody identification. This result may reflect the fact that standard music perception paradigms require the listener to detect manipulations of single notes. As such, these paradigms orient the listener to individual notes, which differs from the way the participants approached the melodies in the current experiment. Other aspects of the design may also have contributed to the observed differences between identification and perception. Whereas the music perception literature often uses forced-choice recognition (old/new) of novel melodies over relatively short retention intervals, the current experiment involved retrieving the title of a well-known melody over a relatively long retention interval. Previous research has shown that manipulations of both retention interval (Dowling & Bartlett, 1981; Edworthy, 1982, 1985) and stimulus familiarity (J. D. Smith et al., 1994) influence performance on music cognition tasks. In any case, the current data suggest that future work in music perception should assess the effects of phrase-level manipulations (see Tillman & Bigand, 1998, for an example of this ap-

proach). In fact, phrase-level manipulations may do a better job of capturing how music is perceived outside of the laboratory.

#### HOW IS THE PHYSICAL STIMULUS RECONCILED WITH ITS CORRESPONDING MENTAL REPRESENTATION?

Although the most consistent predictor of identification performance was serial position, it would be foolish to argue that there is a direct causal relationship between serial position and identification. However, the relationship between identification and serial position is important for understanding how the physical stimulus is matched to its mental representation. First, plotting serial position by identification performance reveals that the majority of melody identifications (77.3%) occurred by SP 7 (Figure 2). This indicates that people are able to identify well-known melodies based on relatively brief exposures. Second, the pattern of identification performance across serial positions traced an inverted U-shape. In fact, beyond some point, additional information was less effective than that which preceded it. One could argue that this pattern was related to differences in the ability of the participants rather than characteristics of the music and/or the attempt to identify the melody. However, this alternative explanation was excluded because the regression analyses that separated the more and less proficient participants produced results that were quite similar to those of the main analyses.

The inverted U-shaped function strengthens the proposed parallel between melody and spoken word identification (Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978). To review, the cohort theory argues of spoken word identification argues relies on temporally ordered, bottom-up processing and claims that a word is identified when the physical input reaches a point that distinguishes it from all other words in the lexicon (Warren & Marslen-Wilson, 1987, 1988). Identification is unlikely to occur before this divergence point, and information that follows the divergence point should be relatively inconsequential. The data from the current experiment conform to these predictions. Performance was relatively low at early SPs, presumably because there was not enough information to distinguish the target melody from other melodies stored in memory. On average, the divergence point for melodies appears to occur between SP 5 and SP 7; this is the region in which identification peaked. The participants who failed to identify the song at this point did so because of noise in the system or a failure to correctly perceive the early segments of the melody. Consequently, information following the divergence point proved to be of little value as performance dropped and remained low across SPs 8–12. This interpretation implies that early segments of the melody are crucial for identification (Schulkind, 2003).



The current data also lead to several other suggestions for future research in the field. For example, the current data always assessed identification starting from the first note in the melody, an approach echoed in research on word identification (Marslen-Wilson, 1987). Different results might be obtained if participants were given notes from different places within the melody. It is also possible that the nature of the stimuli or the musical ability of the participants influenced the results reported here. The use of trained musicians and/or more complex musical forms (e.g., jazz, classical, and even non-Western styles) would provide additional tests for the generality of the current findings. A third important direction to take the current experiment would be to construct novel melodies to provide direct tests of the variables identified as significant predictors. Comparisons of the models for well-known versus novel melodies would have important implications for psychomusicology as discrepancies between different kinds of melodies may provide insight into what makes a melody “good” enough to become a cultural standard (Bartlett, 1993; Bartlett & Dowling, 1988).

In closing, one general methodological issue must be addressed. Although the regression analyses explained a substantial proportion of the variability in identification performance, a critical reader might argue that this leaves more than half the variability in melody identification unexplained. However, several factors—that could not be controlled or measured—limited the predictability of the data. For example, we had no control over exposure, recent listening experiences, or availability, in general. In addition, one would expect melody identification to be a function of both the properties of the melody (which we measured) and the relationship of these properties to those of all other songs the participant knows; a complete assessment of each participant’s knowledge base would not be feasible. This strategy would also require making decisions regarding which properties of the participants’ knowledge base were relevant; this was the express purpose of the current experiment. As well, the participants may have begun to form expectations about the identity of future stimuli based on what they had heard. Finally, there are many strategies and processes in the recognition of sequentially presented stimuli that could interact with the process of recognition (see Schulkind, 2000, 2002). Some of these issues might be addressed in future work using experimentally constructed, novel melodies that would limit the number of stimuli and/or stimulus properties to be considered. We view the current work as a prelude to such research, as a means to guide the selection of research questions in future work. More generally, we acknowledge that the approach taken in the current research may have influenced the results. Whereas we were more interested in the behavior of relatively unsophisticated musicians identifying popular songs, most research in the field tends to be skewed toward

more trained listeners and novel, experimental materials. As was acknowledged throughout the discussion, this orientation might explain some of the discrepancies between our data and previous work in the field.<sup>2</sup>

## References

- Andrews, M. W., Dowling, W. J., Bartlett, J. C., & Halpern, A. P. (1998). Identification of speeded and slowed familiar melodies by younger, middle-aged, and older musicians and nonmusicians. *Psychology & Aging, 13*, 462–471.
- Banich, M. T., & Heller, W. (1998). Evolving perspective on lateralization of function. *Current Directions in Psychological Science, 7*, 1–2.
- Bartlett, J. C. (1993). Tonal structure of melodies. In T. J. Tighe & W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm*. (pp. 19–38). Hillsdale, NJ: Lawrence Erlbaum.
- Bartlett, J. C., & Dowling, W. J. (1988). Scale structure and similarity of melodies. *Music Perception, 5*, 285–314.
- Bartlett, J. C., Halpern, A. R., & Dowling, W. J. (1995). Recognition of familiar and unfamiliar melodies in normal aging and Alzheimer's disease. *Memory & Cognition, 23*, 531–546.
- Bartlett, J. C., & Snelus, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology, 93*, 551–560.
- Bever, T. G., & Chiarello, R. J. (1974). Cerebral dominance in musicians and nonmusicians. *Science, 185*, 537–539.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115–147.
- Boltz, M. (1991). Some structural determinants of melody recall. *Memory & Cognition, 19*, 239–251.
- Boltz, M., Marshburn, E. Jones, M. R., Johnson, W. W. (1985). Serial-pattern structure and temporal-order recognition. *Perception & Psychophysics, 37*, 209–217.
- Butler, D. (1992). *The musician's guide to perception and cognition*. New York: Macmillan, Inc.
- Cohen, A. J. (1990). Understanding musical soundtracks. *Empirical Studies of the Arts, 8*, 111–124.
- Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge, MA: The MIT Press.
- Cuddy, L. L. (1993). Melody comprehension and tonal structure. In T. J. Tighe and W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm*. (pp. 19–38). Hillsdale, NJ: Lawrence Erlbaum.
- Cuddy, L. L., Cohen, A. J. & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance, 7*, 869–883.
- Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics, 28*, 381–389.
- Deutsch, D. (1983). The generation of two isochronous sequences in parallel. *Perception & Psychophysics, 34*, 331–337.
- Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88*, 503–522.
- Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics, 14*, 37–40.
- Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*, 341–354.

2. The authors thank Tamara Rahhal, Mari Jones, Lola Cuddy, and an anonymous reviewer for helpful comments on an earlier draft.

- Dowling, W. J. & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, 1, 30–49.
- Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 49, 524–531.
- Dowling, W. J., Lung, K. M-T., & Herrbold, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics*, 41, 642–656.
- Dyson, M. C., & Watkins, A. J. (1984). A figural approach to the role of melodic contour in melody recognition. *Perception & Psychophysics*, 35, 477–488.
- Edworthy, J. (1982). Pitch and contour in music processing. *Psychomusicology*, 2, 44–46.
- Edworthy, J. (1985). Interval and contour in melody processing. *Music Perception*, 2, 375–388.
- Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Perception & Psychophysics*, 40, 67–68.
- Essens, P. J., & Povel, D-J. (1985). Metrical and non-metrical representations of temporal patterns. *Perception & Psychophysics*, 37, 1–7.
- Gates, A., & Bradshaw, J. L. (1977). The role of cerebral hemispheres in music processing. *Brain and Language*, 4, 403–431.
- Gianetti, L. (1982). *Understanding movies*. Englewood Cliffs, NJ: Prentice Hall.
- Halpern, A. R. (1984) Perception of structure in novel music. *Memory & Cognition*, 12, 163–170.
- Halpern, A. R. (1988). Perceived and imagined tempos of familiar songs. *Music Perception*, 6, 193–202.
- Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, 17, 572–581.
- Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.
- Hébert, S., & Peretz, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory & Cognition*, 25, 518–533.
- Hutchinson, W., & Knopoff, L. (1978). The acoustic component of Western consonance. *Interface*, 7, 1–29.
- Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics*, 41, 621–634.
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and reactions to time. *Psychological Review*, 96, 459–491.
- Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics*, 32, 211–218.
- Jones, M. R., & Ralston, M. T. (1991). Some influences of accent structure on melody recognition. *Memory & Cognition*, 19, 8–20
- Jones, M. R., Summerell L., & Marshburn, E. (1987). Recognizing melodies: A dynamic interpretation. *Quarterly Journal of Experimental Psychology*, 39(A), 89–121.
- Jusczyk, P. W., & Luce, P. A. (2002). Speech perception. In H. Pashler & S. Yantis (Eds.), *Steven's handbook of experimental psychology (3rd ed.)*, Vol. 1: *Sensation and perception* (pp. 493–536). New York: John Wiley & Sons.
- Kameoka, A., & Kuriyagawa, M. (1969). Consonance theory. Part I: Consonance of dyads. *Journal of the Acoustical Society of America*, 45, 1451–1459.
- Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 15, 166–171.
- Kimura, D. (1964). Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, 16, 355–358.
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11(3), 346–374.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, 126, 159–179.

- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 89, 334–368.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, 56, 414–423.
- Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, 58, 927–935.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71–102.
- Marslen-Wilson, W., Moss, H. E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1376–1392.
- Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1–71.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10, 29–63.
- Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 576–585.
- Maylor, E. A. (1991). Recognizing and naming tunes: Memory impairment in the elderly. *Journals of Gerontology*, 46, P207–P217.
- Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, 15, 238–246.
- Monahan, C. B., Kendall, R. A., & Carterette, E. C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. *Perception & Psychophysics*, 41, 576–600.
- Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.
- Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structure in determination of musical phrases. *Journal of Experimental Psychology: Human Perception & Performance*, 13, 116–126.
- Palmer, C., & Krumhansl, C. L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony performance, timing and familiarity. *Perception & Psychophysics*, 41, 505–518.
- Peretz, I., & Kolinsky, R. (1994). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *Quarterly Journal of Experimental Psychology*, 46(A), 301–325.
- Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociation following bilateral lesion of auditory cortex. *Brain*, 117, 1283–1301.
- Peretz, I., & Morais, J. (1987). Analytic processing in the classification of melodies as same or different. *Neuropsychologia*, 25, 645–652.
- Peretz, I., Morais, J., & Bertelson, P. (1987). Shifting ear differences in melody recognition through strategy inducement. *Brain and Cognition*, 6, 202–215.
- Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review*, 77, 481–496.
- Restle, F. & Brown, E. R. (1970). Serial pattern learning. *Journal of Experimental Psychology*, 83, 120–125.
- Robinson, G. M. (1977). Rhythmic organization in speech processing. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 83–91.
- Robinson, G. M., & Solomon, D. J. (1974). Rhythm is processed by the speech hemisphere. *Journal of Experimental Psychology*, 102, 508–511.
- Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, 4, 328–350.
- Schellenberg, E. G., Adachi, M., Purdy, K. T., & McKinnon, M. C. (2002). Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology: General*, 131, 511–537.

- Schulkind, M. D. (1999). Long term memory for temporal structure: Evidence from the identification of well-known and novel songs. *Memory & Cognition*, 27, 896–906.
- Schulkind, M. D. (2000). Perceptual interference decays over short unfilled intervals. *Memory & Cognition*, 28, 949–956.
- Schulkind, M. D. (2002). Feature modulation search: A novel memory search model that extends the perceptual interference effect to musical stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 346–352.
- Schulkind, M. D. (2003). Serial processing in melody identification. Manuscript in preparation.
- Schulkind, M. D., Hennis, L. K., & Rubin, D. C. (1999). Music, emotion and autobiographical memory: They're playing your song. *Memory & Cognition*, 27, 948–955.
- Smith, D. A., & Graesser, A. C. (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. *Memory & Cognition*, 9, 550–559.
- Smith, J. D., Kemler Nelson, D. G., & Appleton, T. (1994). What child is this? What interval was that? Familiar tunes and music perception in novice listeners. *Cognition*, 52, 23–54.
- Stewart, D., & Furse, D. (1986). *Effective television advertising*. Lexington, MA: Heath.
- Tarr, M. J., & Vuong, Q. C. (2002). Visual object recognition. In H. Pashler and S. Yantis (Eds.), *Steven's handbook of experimental psychology (3rd ed.)*, Vol. 1: *Sensation and perception* (pp. 287–314). New York: John Wiley & Sons.
- Tillman, B., & Bigand, E. (1998). Influence of global structure on musical target detection and recognition. *International Journal of Psychology*, 33, 107–122.
- Warren, P., & Marslen-Wilson, W. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception and Psychophysics*, 41, 262–275.
- Warren, P., & Marslen-Wilson, W. (1988). Cues to lexical choice: Discriminating place and voice. *Perception and Psychophysics*, 43, 21–30.
- White, B. W. (1960). Recognition of distorted melodies. *American Journal of Psychology*, 73, 100–107.
- Yalch, R. F. (1991). Memory in a jingle jungle: Music as a mnemonic device in communicating advertising slogans. *Journal of Applied Psychology*, 76, 268–275.

## Appendix

The 34 target songs were: “Amazing Grace,” “America, the Beautiful,” “(Oh my Darling,) Clementine,” “Dixie,” “Do-Re-Mi (Do a Deer),” “Frère Jacques,” “Frosty, the Snowman,” “Happy Birthday to You,” “Home on the Range,” “(If You’re Happy and You Know it) Clap Your Hands,” “Itsy-Bitsy Spider,” “I’ve Been Working on the Railroad,” “Joy to the World,” “Lean on Me,” “London Bridge is Falling Down,” “Mary Had a Little Lamb,” “National Anthem,” “On Top of Old Smokey,” “Pop Goes the Weasel,” “Puff, the Magic Dragon,” “Rock-A-Bye Baby,” “Row, Row, Row Your Boat,” “Rudolph, the Red-Nosed Reindeer,” “Santa Claus is Coming to Town,” “Silent Night,” “Singin’ in the Rain,” “Take Me Out to the Ballgame,” “This Old Man,” “Three Blind Mice,” “Tomorrow,” “Twinkle, Twinkle Little Star,” “When You Wish Upon a Star,” “Yankee Doodle Dandee,” and “Yesterday.”

