Dynamic melody recognition: Distinctiveness and the role of musical expertise

FREYA BAILES

University of Western Sydney, Penrith South, New South Wales, Australia

The hypothesis that melodies are recognized at moments when they exhibit a distinctive musical pattern was tested. In a melody recognition experiment, point-of-recognition (POR) data were gathered from 32 listeners (16 musicians and 16 nonmusicians) judging 120 melodies. A series of models of melody recognition were developed, resulting from a stepwise multiple regression of two classes of information relating to melodic familiarity and melodic distinctiveness. Melodic distinctiveness measures were assembled through statistical analyses of over 15,000 Western themes and melodies. A significant model, explaining 85% of the variance, entered measures primarily of timing distinctiveness and pitch distinctiveness, but excluding familiarity, as predictors of POR. Differences between nonmusician and musician models suggest a processing shift from momentary to accumulated information with increased exposure to music. Supplemental materials for this article may be downloaded from <http://mc.psychonomic-journals.org/content/supplemental>.

A popular 1950s radio show called *Name That Tune* allowed participants to wager on how few notes they would require in order to identify some well-known tune. Experienced listeners are often able to recognize a melody within just a few notes (Dalla Bella, Peretz, & Aronoff, 2003; Schellenberg, Iversen, & McKinnon, 1999; Schulkind, Posner, & Rubin, 2003). The present study provides an examination of how this is accomplished. Specifically, it investigates the factors that contribute to the time course of recognizing or identifying a melody.¹

Cohort theory posits that word recognition is based on the distinction of the particular phoneme sequence of the word in the context of some lexicon. When applied to music, an initial cohort of melodies would be activated on the basis of the first notes of a melodic sequence. Thereafter, members of the initial cohort that do not match the increased information provided by the unfolding melodic sequence are dropped until the correct melody is isolated. A consequence of cohort theory is that a melody's point of recognition (POR) should correlate with an increase in information, or distinctiveness, of the melodic sequence. This is because a distinctive melodic event would serve to eliminate irrelevant melodies from the cohort, leading to isolation and recognition. Some melodies are more distinctive or unusual than others. For example, many melodies begin with an ascending perfect fourth interval, whereas few melodies begin with an ascending tritone. Hence, the first two notes of "Maria" from Leonard Bernstein's *West Side Story* are far more distinctive than the initial notes for "The Farmer in the Dell." Few melodies begin with the same rhythm as "Happy Birthday," whereas many begin

with a series of isochronous durations, including "Frère Jacques."

In addition to their role at retrieval, distinctive events may also be important in melody recognition because of their salience at encoding (see Reder, Paynter, Diana, Ngiam, & Dickison, 2008), with McAuley, Stevens, and Humphreys (2004) speculating that melodies that are not distinctive (or "catchy") are not well attended to. Studies of melody identification (Hébert & Peretz, 1997; White, 1960), recall, and expectancy (Carlsen, 1981) have posited a contribution of musical distinctiveness without directly testing it. Schulkind et al. (2003) examined the musical features that facilitate melody identification and, so, questioned what type of information (e.g., phrase boundaries, melodic interval, musical ornaments) contributes to melody recognition. In their study, 28 participants, who were not selected for musical training, identified 34 songs presented note by note. The relationship of recognition to serial position (i.e., note number) exhibited an inverted U shape, leading the authors to conclude that melodies generally were identified after the presentation of a moderate amount of information (namely 5–7 notes), enough information for unique identification. However, the direction of the supposed influence of distinctive features on melody identification is not always clear in their theoretical predictions. For example, they speculate that relatively large intervals between pitches (which are atypical and therefore distinctive) might facilitate identification, but also that relatively small intervals (which are typical) might facilitate identification by meeting listener expectations. Nevertheless, the work is a useful step to-

ward a direct analysis of the way in which typical versus atypical musical properties facilitate melody recognition.

Despite a strong theoretical framework and our demonstrated ability to identify speech from minimal information (Schellenberg et al., 1999), empirical evidence for the positive role of atypical or distinctive features in memory for music is lacking. For example, Schmuckler (1997) asked participants first to rate the degree of expectancy of different melodic endings and subsequently to perform a surprise melody recognition task. Participants better recognized those melodies rated as “expected,” suggesting the presence of a facilitating effect in short-term memory from the *typicality* of their construction. Halpern, Müllensiefen, and Wiggins (2008) reported little success in attempts to incorporate the distinctiveness of a musical feature when melody recognition was modeled.

In contrast to visual recognition (such as recognizing faces), music-related recognition more obviously involves the progressive buildup of information over time. Nevertheless, very few dynamic models of melody recognition have been developed. Rare exceptions are studies by Dalla Bella et al. (2003) and Schulkind et al. (2003), which made use of a gating paradigm, a method traditionally used for research in auditory word recognition. This method presents segments of auditory information in which successive presentations increase the duration of the segment. Applied to melody recognition, segments are systematically increased in the number of notes for each presentation of a stimulus melody. In Dalla Bella et al.’s study, melody recognition was assessed after each segment, either through familiarity judgments (Experiment 1) or by asking participants to sing the melodic continuation (Experiment 2). As in Schulkind et al. (2003), listeners were able to identify the melody within five to seven notes.

There are two ways to conceptualize the information necessary to trigger melody recognition: First, information might accumulate (*accumulated information*) through the onset of successive notes until the stimulus becomes wholly distinctive, at which point, if the melody is familiar, the listener will indicate recognition (Schulkind et al., 2003). Second, a highly distinctive moment (*momentary information*), such as an unusual pitch interval, might be sufficient to precipitate recognition (see Schulkind et al., 2003; Stevens & Latimer, 1997). In general, accumulated information could be expected to coincide with momentary information, since a certain amount of information is required to trigger melody recognition, be this through the presentation of one note or of many notes. However, it is conceivable that an event is so highly distinctive, or is even unique, that it might render any preceding notes and their accumulated information redundant. Also, information might not accumulate in a linear fashion, with the forgetting of events that happened a while before lending more weight to momentary information. For this reason, in the present study, both momentary and accumulated information are examined as measures of distinctive events.

What is and what is not distinctive for an individual must depend on those patterns to which the individual has been exposed. Increased exposure to music, as occurs through musical training, might “serve to neutralize those

unique expectancies which would be a residual of early musical experience” (Carlsen, 1981, p. 20). Alternatively, musicians might simply increase their cohort without distorting the proportion of distinctive events it contains. Evidence for expertise effects in memory for melody is mixed. Peynircioğlu, Tekcan, Wagner, Baxter, and Shaffer (1998) found that musicians with more than 5 years of training and those who listened to more than 20 hours of music a week were no better at recalling well-known melodies (e.g., from a title) than were other participants. Dalla Bella et al. (2003) compared melody recognition for those with and without musical training. Group differences were found and were explained in relation to cohort theory: The authors argued that the musicians in their study knew more pieces of music than did those without musical training. This would represent a larger initial cohort and might explain a later isolation point for musicians than for nonmusicians, which is the note at which participants indicate that they know the melody. However, although musicians were later in isolating the melody from the cohort than were nonmusicians, they were earlier than nonmusicians in reaching a *recognition point*, the note at which total confidence is achieved. Cohort theory would predict such a result if the musicians are more familiar with the melodies than are the nonmusicians, since it argues that the level of activation of items in the initial cohort varies as a function of frequency. In the *word-frequency effect*, high-frequency words are generally recognized more rapidly than low-frequency words. Translated to music, a highly familiar melody may elicit a higher initial level of activation than would less familiar melodies, so that less activation is required to reach the threshold for isolation.

It is appropriate to question what kind of information structures melody and, hence, contributes to its recognition. Evidence has been found in favor of roles for pitch, pitch interval (Carlsen, 1981; Dowling & Bartlett, 1981; White, 1960), scale degree (Pearce & Wiggins, 2006), melodic contour (Davies & Yelland, 1977; Dowling & Bartlett, 1981; Dowling & Fujitani, 1971; Idson & Massaro, 1978; Schulkind et al., 2003), duration, and phrasing or grouping (Dalla Bella et al., 2003; Schulkind et al., 2003). Cuddy, Cohen, and Miller (1979) wrote that the ability to apply a tonal framework to a sequence is a critical determinant to recognition (see also Krumhansl & Kessler, 1982; Schulkind et al., 2003). In song, lyrics provide an additional layer of information. Finally, meter (patterns of beat, recurring in a hierarchical relationship) is a potentially informative structural element. For instance, Schulkind et al. found that melody recognition was significantly better on notes falling on the first beat of the measure than on any other beat of the measure.

Although it is generally acknowledged that a combination of pitch and rhythm leads to optimal melody recognition (Hébert & Peretz, 1997; Pearce & Wiggins, 2006; Schulkind et al., 2003), a handful of experiments suggest that pitch may be more important than rhythm in melody recognition (e.g., Dalla Bella et al., 2003; Demorest & Kim, 2004; Hébert & Peretz, 1997). Interestingly, in relation to the present study, an explanation might stem from what Hébert and Peretz describe as the lesser encoding dis-

tinctiveness of rhythmic structure relative to pitch structure in Western musical tradition. "The music of our culture is simply more elaborate in pitch than in duration, and it uses pitch categories that are fixed to certain frequencies, whereas durational categories are more flexibly defined" (Hébert & Peretz, 1997, p. 530). "It just happens to be characteristic of most music in our culture that rhythmic patterns are stereotyped and afford a less adequate basis for identification than melody" (White, 1960, p. 103).

The present article argues that melodies are recognized at moments in which they exhibit a unique or distinctive musical pattern, with the prediction that distinctive pitch events would be more associated with the POR than would distinctive temporal events, but that both would contribute to successful melody recognition. However, the evidence to date does not clearly support a positive association between melody distinction and enhanced recognition. Schulkind et al. (2003) explored dynamic melody recognition, but they did not take the subsequent step to associate the information specified by the structural elements at each serial position with the POR. Although Peynircioğlu et al. (1998) examined differences between musicians and nonmusicians with respect to melody recognition, their experimental tasks did not tap the dynamic process of recognizing a melody, wherein differences in distinctiveness might play a role.

Melody recognition is clearly dynamic, and one aim of the present study was to integrate the temporal approach of Dalla Bella et al. (2003) with a more detailed focus on distinctive structure than provided by the initial ideas advanced by Schulkind et al. (2003) and the static analyses by Halpern et al. (2008) that summarized across time. A large corpus of melodies is used to calculate the frequency of occurrence and, hence, information values of different melodic elements. These are then used to model the time course of melody recognition as measured in a perceptual experiment comparing the points of melody recognition² for musician and nonmusician participants.

The present study additionally examines the interaction among different types of information during melody recognition, namely pitch and timing, as well as momentary and accumulated information. There are many different aspects of pitch that might be represented. Information about a pitch in isolation is not salient, except for listeners with absolute pitch (the ability to identify a pitch in the absence of an external referent). More promising as a melodic feature rich in information is scale degree, since this allows two melodies to be recognized as identical, even when they are transposed. Accordingly, scale-degree information is studied and so is melodic interval. This can comprise two parts: magnitude and contour direction (ascending or descending) (Schulkind et al., 2003; Stevens & Latimer, 1997). The present study incorporates both parts, for a measure of interval magnitude and direction (comparable to "Int*Dir" in Schulkind et al., 2003).

As for timing information, notes of long duration may act as temporal accents, thereby attracting attention and facilitating melody recognition (Schulkind et al., 2003). Moreover, there is increasing evidence of the importance of durational contrast in speech and music perception (Huron

& Ollen, 2003). Accordingly, the present study measures the relative duration of three subsequent interonset intervals (similar to "Rel-Dur" in Schulkind et al., 2003). In other words, a duration longer than both the previous and subsequent durations would be coded "longer–shorter." Meter is also taken into account. As with duration, it is coded as a relative measure of three subsequent events. For instance, a note in a metric position that is stronger (than that of the previous note) followed by a note of weaker metric position would be coded "stronger–weaker."

By testing the accuracy of different measures of pitch and timing information to predict the moment at which listeners recognize melodies, this research can directly question the importance of distinctive events in retrieval for listeners with and without musical training.

Stimulus Selection

Because familiarity and distinctiveness are at least conceptually different, stimuli for the melody recognition experiment were selected from two groups of melodies, each intended to maximize the variance for familiarity and distinctiveness, respectively.

Familiarity. For this subsample, the aim was to select three levels of melodic familiarity: extremely, moderately, and much less familiar. In order to help select these musical materials, 13 respondents were recruited for a pretest survey: 7 music majors and 6 nonmusic majors, drawn from the same pools used in the main experiment. Each pretest participant heard the first two phrases of 160 melodies, which were drawn from the same sources that were used for the sample of melodic stimuli employed in a subsequent recognition experiment. Specifically, the sources comprised 6,174 Germanic folk melodies from Schaffrath's (1995) Essen Folksong Collection database, 9,788 classical themes (essentially tonal or modal) from Barlow and Morgenstern's (1948) *Dictionary of Musical Themes*, and 107 traditional melodies known to American listeners and taken from four "fake book" sources. From these sources, 160 melodies with an average length of 45.6 notes were selected. Participants listened in three mixed groups of musicians and nonmusicians, with the melodies presented in different random orders. They rated their familiarity with each melody on a 7-point Likert scale that ranged from *extremely familiar* (e.g., "Farmer in the Dell," with a mean familiarity score of 7.0) to *extremely unfamiliar* (e.g., the first theme from Gian Francesco Malipiero's *Impressioni dal vero*, with a mean familiarity score of 1.2). From the original 160 melodies, 60 were selected randomly and conceptually divided into three 20-melody sets: one from the upper third (*familiar*, range 6.5–7.0), one from the lower third (*unfamiliar*, range 2.54–3.01), and one from the remaining third (*moderately familiar*, range 3.45–6.12).³

Distinctiveness. The melodies selected represented three levels of overall distinctiveness: highly, moderately, and least distinctive. From the original 160 pretest melodies, 60 already had been selected for the familiarity stimuli. This left 100 melodies for further consideration and for which a distinctiveness value was calculated.

Distinctiveness measures for each note were computed by using the following method: First, the probability of a

Table 1
Scale-Degree Information in Bits for Major and Minor Keys

Scale Degree	Major Key Information (in Bits)	Minor Key Information (in Bits)
1	2.46741	2.36813
1+	8.71624	10.1067
1–	11.2996	10.6917
2	2.76651	2.83917
2+	9.13073	13.499
2–	9.82328	7.50278
3	2.45858	2.63612
3+	10.4012	7.08116
3–	7.18182	14.499
4	3.25139	3.03107
4+	6.92843	6.8991
4–	11.0398	10.3026
5	2.27553	2.28662
5+	9.01341	12.499
5–	10.4385	10.3026
6	3.60456	4.2194
6+	10.444	5.94187
6–	8.25012	11.8361
7	4.04147	4.83546
7+	10.5744	9.03959
7–	7.25964	4.42398

Note—+, augmented; –, diminished.

given event, such as the interval of a rising perfect fifth, was calculated using Humdrum (Huron, 1994), which computes the frequency of occurrence, and, hence, the probability, of events in a defined musical corpus. Here, probability values were obtained from a statistical analysis of notes occurring in the same database of German folk melodies, classical themes, and traditional melodies from which the stimuli were selected (i.e., 16,069 melodies). Next, the corresponding information value for an event was calculated. This value is the base-2 logarithm of the probability of occurrence and typically is expressed in binary digits, or *bits*.

Five measures were used in calculating an overall distinctiveness value. Three measures pertain to pitch-related information (i.e., interval, scale degree, and scale-degree dyad), and two measures pertain to timing-related information (i.e., metric stress [strong/weak] and rhythm). See Huron (2006).

With regard to interval information, the probabilities of occurrence for melodic intervals from the database of musical scores were measured. For example, the most common interval is the descending major second ($p = .195$), and the least common is the descending augmented fourth ($p = .0006$). The corresponding information value (in bits) was then calculated for each interval (refer to the online supplemental materials for interval information values). For example, the descending major second, with a probability of occurrence of .195 (collapsed across major and minor key contexts) has an information value of 2.35845 bits.

With regard to scale-degree information, the probabilities of occurrence for various scale degrees were measured, and the probabilities for each note in the melody were characterized. The corresponding information value was calculated for each scale degree (see Table 1).

With regard to scale-degree-dyad information, the probabilities of occurrence for various pairs of successive notes were measured and the probabilities for successive note pairs in the melody were characterized. The corresponding information value was calculated for each scale-degree dyad (refer to the online supplemental materials for scale-degree-dyad information values).

In the case of metric-stress information, the relative metric stress of successive pairs of notes was characterized. For example, in 4/4 meter, the downbeat and half-measure beats were deemed the strongest and second-strongest onset moments. The second and fourth beats were deemed less strong. The half-beat metrical positions were weaker yet, and the quarter-beat positions were the weakest. Using such a metrical hierarchy, each note was determined to be stronger, weaker, or of the same strength as the previous note. The probabilities of occurrence for various metric-stress dyads were measured, and each note pair in the target melody was characterized by using this probability. The corresponding information value was calculated for each metric-stress dyad (see Table 2).

In the case of rhythmic information, the relative duration of successive pairs of notes was characterized (see Table 3). Each note was determined to be longer, shorter, or of the same duration as the previous note. Rests were amalgamated with the preceding note, so a quarter note followed by a quarter rest was deemed equivalent to a half-note duration.

Rather than attempt to characterize the melodic distinctiveness of an entire melody, we focused on the start of the melody. In a pilot experiment, POR data were gathered for 13 melodies. Table 4 shows the recognition percentage for different points in the melody for those melodies that were identified by participants.

Notice that, after roughly 8 notes, recognition scores do not increase appreciably with increased length of the melodic incipit. Clearly, and in line with the literature, the first 5 or 6 notes are especially valuable in melody identification. Consequently, to obtain a global distinctiveness value, the measures were weighted according to the serial position in the melody. Hence, the 3rd note was weighted as 13.6, the 4th was weighted as 29.6, and so forth. The calculations included only the first 10 notes of the melody.

The melody's final distinctiveness rating was calculated as follows: For each note, all five information values were averaged. This average was then weighted by

Table 2
Information in Bits Associated
With Different Patterns of Metric Stress

Metric Stress Pattern	Information (in Bits)
Same–same	8.20238
Same–stronger	4.48003
Same–weaker	5.77787
Stronger–same	5.91717
Stronger–stronger	4.96893
Stronger–weaker	1.33506
Weaker–same	4.41636
Weaker–stronger	1.40765
Weaker–weaker	3.94435

Table 3
Information in Bits Associated
With Different Patterns of Relative Event Duration

Relative Durations	Information (in Bits)
Longer-longer	5.46414
Longer-same	4.09479
Longer-shorter	2.48366
Same-longer	2.74432
Same-same	1.68473
Same-shorter	4.80715
Shorter-longer	3.27171
Shorter-same	3.09634
Shorter-shorter	5.40837

the serial-position weighting factor to give a weighted information value. Then the weighted information values for the first 10 notes were averaged to give a distinctiveness rating per melody.

From the 100 melodies, three groups of 20 melodies were selected randomly as follows: The first group came from the upper third of the distinctiveness means (*highly distinctive*, 4.19–4.83 bits), the second came from the lower third (*not distinctive*, 3.24–3.76 bits), and the third came from the remaining third (*moderately distinctive*, 3.80–4.15 bits).

By way of summary, we selected 120 stimulus melodies (listed in the supplemental materials), with 60 melodies in each of two classes (familiarity and distinctiveness).⁴ Within each class, there were three levels of familiarity or distinctiveness: high, medium, and low. Consequently, we used six categories of melodic stimuli, which included instrumental melodies with no associated lyrics (72, including unfamiliar German folk songs), as well as songs (48, excluding unfamiliar German folk songs). Song lyrics and melody titles were not presented to participants.

MELODY RECOGNITION EXPERIMENT

To assemble POR data, a melody recognition study was conducted in which listeners were exposed to melodic fragments and asked to indicate the point at which they recognized a melody.

Method

Participants. Of the 32 experimental participants (17 males, 15 females; mean age = 26.9 years) recruited from members of the Ohio State University community, 16 were musicians (primarily second-year music majors who participated in the experiment for academic credit) and 16 were nonmusicians (students recruited from several non-music classes). The nonmusicians received payment for their participation, and all but 5 of the participants were born in the United States. None had participated in the stimulus-selection pilot testing.

Stimuli. Each participant heard 5 melodies from each of the six categories of melodic stimuli. The order of presentation was randomized. In selecting the melodies for a given participant, the melodies were selected randomly without replacement; subsequent participants continued sampling from the remaining stimuli in the melody pool. This meant that, although each participant heard only 30 melodies, each group of 4 participants collectively provided data for all 120 stimulus melodies. With 32 participants in total, data were collected for eight complete 4-participant sets of all 120 melodies.

Procedure. In order to establish the POR for each melody for each participant, participants heard the melodies using a gating

paradigm. The experimenter played the first note on an electronic keyboard (set to a piano sound). The experimenter then played the first 2 notes, then the first 3 notes, and so forth. Participants were asked to indicate when (and if) they recognized the tune. Participants could indicate recognition either by naming the tune or by humming or singing the melodic continuation. Unlike Schulkind et al. (2003), the present study included stimuli other than songs with well-known lyrics. Thus it was important to allow for the possibility that participants recognized the melody (i.e., identified the note sequence), even when they could provide no verbal label to indicate recognition. The current note was deemed to be the POR, if, in the opinion of the experimenter (a trained musician), the participant had successfully identified the tune name or sung an adequate continuation of the melody in question.⁵ Participants were told that their goal was to identify the tune as soon as possible. Feedback was provided after each guess. The trial was stopped as soon as the melody had been identified or after unsuccessful presentation of the first 10 notes. Using this procedure, recognition data for 960 cases were collected.

Following the presentation of all 30 trials, the second part of the experiment involved revisiting all of the melodies to collect individual familiarity data. Participants were asked to identify how familiar they were with each melody on a 7-point scale, where 1 = *not familiar at all with the melody* and 7 = *very familiar with the melody*.

Results

Listeners failed to recognize melodies on 628 trials. Of the 332 melodies that were recognized,⁶ the mean POR was 5.4 notes with a mode of 4. Figure 1 shows this distribution.

As was the case in the study by Schulkind et al. (2003), a potential confound was that participants would be biased by their relative success or failure in the first task when rendering familiarity judgments. If the familiarity ratings are not influenced by such a bias, one would expect a very high correlation between the familiarity judgments for the various stimuli and the familiarity judgments offered by participants in the pretest. To that end, a correlation between the experimental participants' familiarity ratings and the randomly matched pretest participants' familiarity ratings was made. The ratings were broadly similar ($r = .68$).

Returning to the melody recognition data, a negative correlation between POR and familiarity suggests that early recognition tends to occur when the melody is rated as highly familiar [$r(330) = -.77, p < .001$].

A comparison of musician and nonmusician responses shows differences in POR, with musicians returning an

Table 4
Cumulative Melody Recognition From Notes 1 to 10

No. of Initial Notes	Cumulative % Recognition	% Increase
1	0	—
2	16.0	16.0
3	29.6	13.6
4	59.2	29.6
5	79.2	20.0
6	87.2	8.0
7	92.8	5.6
8	94.4	1.6
9	95.2	0.8
10	96.8	1.6

Figure 1. Distribution of point-of-recognition responses.

earlier mean POR of 5.1 ($SD = 1.9$) than did nonmusicians ($M = 5.6$, $SD = 2.1$). However, musicians reported greater familiarity with the stimuli ($M = 4.2$, $SD = 2.6$) than did nonmusicians ($M = 3.6$, $SD = 2.5$).

It could be argued that a mean POR of 5.4 notes occurs as a result of having heard (ever-increasing) portions of the melody five or more times. In other words, perhaps this has little to do with the moment in the melody at which recognition is achieved, but instead reflects that melodies need to be heard a certain number of times for them to be recognized. If this were the case, then structural information would not be an important predictor of POR. In order to test this explanation and to determine the best predictors of melody POR, a multiple regression analysis of corresponding familiarity and distinctiveness values was run.

MODELING

Predictor Variables

Given the demonstrated importance of familiarity, the first predictor variable was subjective familiarity with the melody. The factors of key theoretical interest for their predictive value were measures of distinctiveness. Ten information theoretic measures were divided into two groups of five: *Momentary measures* pertain to information afforded by the event itself, and *accumulated measures* pertain to the total accumulated information up to the current point in the melody. Conceptually, momentary measures might be important if particular sorts of events are especially telling and so precipitate recognition of the tune. Accumulated measures might be important if all of the events leading up to some moment contribute to recognition. For example, recognition may tend to occur when three bits of information have accumulated.

Momentary scale-degree distinctiveness is indicated by the number of bits corresponding to the probability of occurrence of a particular scale degree. This value is high for improbable pitches (such as the raised tonic) and low for more probable ones (such as the dominant). *Momentary interval distinctiveness* is indicated by the number of bits corresponding to the probability of occurrence of particular intervals. This value is high for improbable in-

tervals (such as the augmented fourth) and low for more probable intervals (such as the descending major second). *Momentary scale-degree-dyad distinctiveness* is indicated by the number of bits corresponding to successive pairs of pitches. This value is higher for improbable successions (such as the subtonic pitch followed by the leading tone) and low for more probable successions (such as the leading tone followed by the tonic).

Momentary metric-stress distinctiveness is indicated by the number of bits corresponding to the relative metric stress of triples of notes (calculated from a moving window for each successive note in the melody). The onsets of pairs of notes can be characterized according to their relative placement in the metric hierarchy. For example, a downbeat followed by the second beat would represent a weaker metric relationship (i.e., strong to weak). A pick-up to the downbeat of the next measure would represent a stronger metric relationship (weak to strong). Two successive downbeats would represent a same metric relationship. For three-note groups, nine metric relationships are possible: same-same, same-stronger, same-weaker, stronger-same, weaker-same, stronger-stronger, stronger-weaker, weaker-stronger, and weaker-weaker. The momentary metric-stress distinctiveness is the number of bits corresponding to the probability of occurrence of each of these nine metric relationships.

Momentary rhythmic distinctiveness is indicated by the number of bits corresponding to the relative duration of notes arranged in three-note sequences (calculated from a moving window for each successive note in the melody). Pairs of notes can be characterized according to their relative length (e.g., long-short, short-long). In the case of three successive notes, nine relationships are possible: same-same, same-shorter, same-longer, shorter-same, longer-same, shorter-shorter, shorter-longer, longer-shorter, and longer-longer. Momentary rhythmic distinctiveness is indicated by the number of bits corresponding to the probability of occurrence of each of these nine durational relationships.

In parallel with the five measures above, a series of five accumulated measures was created. For example, the *accumulated scale-degree distinctiveness* represents that total scale-degree distinctiveness measure for all of the notes up to and including the current note in the melody. Rather than representing the information of the particular note, this measure represents the accumulated information for the particular note plus all preceding notes. Such accumulated measures were calculated for each of the representations described above, including interval distinctiveness, scale-degree-dyad distinctiveness, metric-stress distinctiveness, and rhythmic distinctiveness.

These 10 information theoretic measures were produced for each note of each stimulus, derived from the same corpus as that used to calculate weighted aggregate distinctiveness for each melody in the selection of the stimuli (described in the Stimulus Selection section above). However, it should be noted that, in the present analysis, distinctiveness measures did not undergo the serial-position weighting used in the initial categorization of melody distinctiveness.

What is of importance is not the absolute information value for a given note, rather its relative value in the context of a musical sequence. The first transformation of the data was to calculate the increase or decrease in information represented by the onset of each new note in each stimulus (for each of the 10 information theoretic measures). This is akin to a first-order transformation. For example, the scale degree 1 followed by the scale degree 1 would represent no change in information, and consequently receive a difference value of 0. There may be many increases in information of a certain parameter in the space of a number of notes, but it was predicted that the greatest increase in information up to the note in question would be the best predictor of POR. Consequently, for each POR data point, the sequence up to and including the note of POR was assessed for the note of greatest increase in information, for each of the 10 information theoretic measures. These notes of greatest information (for both momentary and accumulated measures) became predictor variables for the multiple regression analysis outlined below.

Figure 2 presents an example of the calculation of momentary and accumulated scale-degree-dyad distinctiveness for “Swing Low, Sweet Chariot.” In the melody recognition experiment, this melody elicited two different PORs, namely Notes 6 and 7. It can be seen that in terms of scale-degree dyad, the most informative note in momentary and accumulated measures is Note 6 (marked in bold). Since the addition of Note 7 does not provide any more informative event, Note 6 is the predictor value. No scale-degree-dyad information is available for Note 1, since it relies on the onset of a second note for its calculation. Such undefined feature information could not contribute to distinctiveness calculations for the very beginning of melodies.

Criterion Variable

The goal of the present study was to determine which, if any, measures of familiarity and distinctiveness predict point of melody recognition. Accordingly, serial position (i.e., note number) of the POR on a recognized trial was taken as the criterion variable. For example, if Listener A identified “Swing Low, Sweet Chariot” at the seventh note, then “7” was entered as the criterion variable in the multiple regression analysis described below. This resulted in from 0 to 8 entries per melody (i.e., between 0 and 8 of the participants exposed to a particular stimulus melody

recognized it), for a total of 332 entries. Individual trial data rather than averaged POR per melody were preferred to obtain the number of observations required for the multiple regression analysis (Tabachnick & Fidell, 1996).

Given the positive skew of POR data (see Figure 1), these were normalized using a log transform before being used as the criterion variable.

Multiple Regression Analyses

Possible problems may occur in using variables determined by time series (i.e., on a serial, note-by-note basis) in the context of a multiple regression analysis. To rule out such time dependencies, all variables to be entered in the regression were checked for the presence of autocorrelation, but none was found. Regression analyses were conducted in SPSS (Release 17.0 for Macintosh). Data were screened for outliers, resulting in the omission of 2 out of 332 cases. Tolerance values were high, indicating low levels of collinearity.

The stepwise method yielded a significant model: $F(7,322) = 264.3, p < .001$. The model explains 84.9% of the variance (adjusted $R^2 = .849$). Table 5 gives information for the predictor variables that are included in the model. The beta coefficients for these variables indicate that POR was associated with an accumulation of rhythmic, scale-degree-dyad, scale-degree, and metric-stress information, and distinctive patterns of metric stress, distinctive intervals, and scale-degree dyad. Note that familiarity, momentary scale degree, momentary rhythm, and accumulated interval were excluded.

Additional analyses were conducted to compare models for musician and nonmusician data. Since musicians had a higher recognition rate than did nonmusicians, a total of 180 data points were available for the musician group and 150 were available for the nonmusician group. Regression analyses for these two groups used the same procedure as the main analysis.

For the musician data, six of the seven variables from the overall model entered into the regression equation: accumulated rhythm, momentary metric stress, accumulated scale-degree dyad, momentary interval, accumulated metric stress, and accumulated scale degree [$F(6,173) = 201.2, p < .001$]. As can be seen in Table 5, the model explains 87.0% of the variance (adjusted $R^2 = .870$).

For the nonmusician data, the same six variables entered into the regression equation as for the musician data,

	2	3	4	5	6	7	Predictor
Momentary scale-degree dyad	5.4629	5.3129	5.4629	4.6150	6.6183	4.6065	
Change in momentary scale-degree dyad	N/A	-0.1500	+0.1500	-0.8479	+2.0033	-2.0118	Note 6
Accumulated scale-degree dyad	5.4629	10.7758	16.2387	20.8537	27.4720	32.0786	
Change in accumulated scale-degree dyad	N/A	+5.3129	+5.4629	+4.6150	+6.6183	+4.6065	Note 6

Figure 2. Scale-degree-dyad distinctiveness for beginning of “Swing Low, Sweet Chariot.”

Table 5
Predictor Variables Included in the Models

Measure	Overall	Musicians	Nonmusicians
Accumulated rhythm	.640	.674	.041
Momentary metric stress	.109	.115	.614
Accumulated scale-degree dyad	.061	.054	.113
Momentary interval	.022	.020	.020
Accumulated scale degree	.010	.005	.024
Accumulated metric stress	.008	.007	.006
Momentary scale-degree dyad	.003		
Variance Explained			
R^2	.852	.875	.818
Adjusted R^2	.849	.870	.811

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.

but in a different order: momentary metric stress, accumulated scale-degree dyad, accumulated rhythm, accumulated scale degree, momentary interval, and accumulated metric stress [$F(6,143) = 107.3, p < .001$]. The model explains 81.1% of the variance (adjusted $R^2 = .811$).

DISCUSSION

The present study was an investigation of the factors that contribute to the time course of melody recognition for musicians and nonmusicians. Multiple regression analyses generated powerful models that predict the POR of a melody on the basis of the information content of specific melodic features. Using a combination of six or seven distinctiveness factors, these equations account for between 81% and 87% of the variance. In other words, as hypothesized, there is a significant relationship between the moment at which listeners recognized melodies in the melody recognition experiment and the presentation of the most informative note in a given parameter.

Early melody recognition correlates with high familiarity ratings, and, as has been found in previous research (Dalla Bella et al., 2003), musically trained participants reported higher levels of familiarity with the stimuli than did nonmusicians. Nevertheless, both overall and group-specific regression analyses excluded familiarity as a factor. This could relate to the ultimately limited variance in familiarity ratings for recognized melodies (overall ratings were dichotomous in distribution). Although a minimal level of familiarity is necessary to recognize a melody, it seems that subjective familiarity ratings do not significantly predict POR. Koriati and Levy-Sadot (2001) made a relevant argument in their study of “feeling of knowing” phenomena: “It is primarily when familiarity is strong enough to drive memory search that the amount of accessible information comes to play a role” (p. 37). The contribution of familiarity in dynamic melody recognition appears to be one of a necessary level of initial activation upon which other factors can build. The present study has demonstrated that these other factors are in large part a combination of types of melodic distinctiveness.

The regression models allow us to examine in more detail the parameters that contribute to recognition. Both time

and pitch measures were tested, and both were included in the final model. Accumulated rhythm and momentary metric stress carried the most weight. Accumulated scale-degree dyad also featured significantly, followed in significance by momentary interval. Accumulated scale degree entered all equations.

The contributions of both time and pitch information are in line with literature supporting an optimal reliance on both in melody recognition (Hébert & Peretz, 1997). However, the fact that accumulated rhythm received such a high loading in the models was unexpected and provides evidence that counters Hébert and Peretz’s theory, whereby rhythm supposedly lacks distinctiveness at encoding and thus has a lesser contribution to melody recognition than do aspects of pitch. Schulkind et al. (2003) were similarly surprised by their empirical finding of a greater contribution to melody recognition from temporal than from pitch factors. They suggest that duration may facilitate melody recognition, with long notes affording time for the listener to analyze what has just been heard, and that rhythm may serve to guide attention. Research has shown that listeners can be more sensitive to rhythmic accents than they are to accents arising from a change in pitch contour (Pfordresher, 2003). Future work might use the same measures of distinctiveness employed in the present study to model participants’ recognition of the stimuli stripped of pitch content or rhythmic content. The results presented here strongly suggest that a rhythmic model for rhythm-only stimuli would outperform a pitch counterpart, but such future work should also contribute by revealing the extent to which the combined influence of pitch and temporal information is additive or multiplicative. Another potentially fruitful approach would be to construct distinctiveness measures based on the probability of occurrence of interacting pitch and timing parameters (e.g., a long rising major third followed by a short falling major second).

It may seem strange that accumulated scale-degree information has a role to play. After all, scale-degree information seems to provide information about only a single pitch. However, recall that, by situating this pitch within a scale, a tonal function is inferred, and an accumulation of such information is indicative of not just melodic but also harmonic distinctiveness. Schulkind et al. (2003) found that tonal function was not a significant predictor of melody recognition and suggested that this might be attributable to a song corpus containing few nondiatonic tones and, hence, little distinction. In contrast, the present study comprised melodies that range from largely diatonic (e.g., “He’s Got the Whole World in His Hands”) to much more chromatic (e.g., Mendelssohn’s “Wedding March”), ranging from low (9.23 bits by Note 4) to high (17.04 bits by Note 4) distinctiveness.

Both accumulated and momentary information measures feature in the regression models, although differences appear between musicians and nonmusicians. Musicians seem to rely more on an accumulation of information about rhythm than do the nonmusicians, for whom the momentary measure of metric stress dominates. Further investigation is needed into the differential contributions

of momentary and accumulated information relative to musical experience. With respect to melody recognition, Schulkind et al. (2003) were interested in trying to separate what they term *holistic properties* from analytic properties of music. Although Schulkind et al. did not study musical expertise (but suggested that doing so would be worthwhile), and although there is no direct mapping of the concepts of analytic and holistic properties to the present distinction between momentary and accumulated information, a common factor is the potential difference between the processing of musical features in isolation or in relation to others. The former is assumed to be context free, whereas the latter relies on a contextual memory, perhaps developed by extensive musical training. Accordingly, it might be that listeners without musical training tend to rely on momentary, distinctive features to recognize melody, whereas musicians draw on the information in a melodic sequence in a cumulative fashion. Experts in a given domain learn to detect patterns, and in the case of musical notes, these extend through time. The present results may reflect a propensity for musicians to group and process longer chains of notes than nonmusicians do.

In the present study, the occurrence of distinctive notes strongly predicted POR. That is not to say that POR can be accounted for entirely by the measures of distinctiveness applied here. A focus on first-order information has neglected any direct study of larger scale musical structures, such as phrase (see Schulkind et al., 2003). An approach able to integrate measures of note-by-note information with higher-order structure might improve the already strong predictive power of the models. For instance, in measures of accumulated information, it could be argued that, having reached the end of a phrase, the subsequent note should not simply be appended. Rather, the start of a new phrase represents a new structure and may be perceived almost in the context of a blank slate rather than as a musical continuation. Some form of weighting might be necessary to account for this effect. However, although phrase and grouping are undoubtedly important cues to melody recognition, this occurs because of the coincidence of the musical features measured here. In other words, a certain amount of larger scale structure is reflected in the information values assigned to the relevant musical features.

A further limitation of the present study concerns the impossibility of tailoring measures of distinctiveness to each listener's individual musical lexicon (Schulkind et al., 2003). There is an important difference between participants' knowledge of specific musical works (i.e., *veridical knowledge*) and their awareness of general musical features characteristic of Western melodies (*schematic knowledge*) (Bharucha, 1987, 1999). In interpreting event-related probabilities of a given corpus, one might regard them as summary statistics that capture something akin to schematic melodic knowledge. However, the melody recognition experiment is quintessentially a task that focuses on veridical knowledge. Although schematic information might play some role in highlighting the distinctiveness of one tone sequence from another, it is possible that veridical knowledge is independent of schematic distinctiveness.

Evidence was found that pitch and timing distinctiveness strongly predicts melody recognition using a mix of song and instrumental melody, but future work should carefully assess any impact of associated lyrics as an integral part of the musical structure. Song lyrics reinforce the rhythmic and metric properties of melody and might partially underpin the strength of rhythmic information as a predictor in the present study. Even the association of a verbal label such as a title might be expected to affect its encoding (McAuley et al., 2004). Indeed, the personal associations that an individual may have in relation to a melody have not been examined in the present research. A more encompassing model of dynamic melody recognition should account for the effects of any deeper melody encoding that may arise from timbral, verbal, and strong affective association.

Finally, translating cohort theory from linguistics to music has entailed important modifications. The gating paradigm originally increased a phonetic segment by a time constant rather than manipulating a variable note length (Dalla Bella et al., 2003). Focusing on notes or events leaves the potential impact of rate of presentation, or tempo, unknown. In a post hoc analysis of the relationship between POR and music tempo, melodies from the highly and moderately familiar categories were sorted with respect to those that could clearly be described as either slow or fast. Eliminating all data for trials in which melodies were not recognized, the POR for four relatively fast melodies ("Can-Can," "Simple Gifts," "Reveille," and "Camptown Races," $N = 26$) and four relatively slow melodies ("Taps," "The Star-Spangled Banner," and "Promenade" and "The Great Gate of Kiev" from Modest Mussorgsky's *Pictures at an Exhibition*, $N = 23$) were examined. Indeed, the mean POR for the selected fast melodies was 5.19 notes, and the mean POR for the selected slow melodies was 4.26 notes. This might suggest that POR relates more meaningfully to a time unit than to a number of notes, with such an approach improving the model's predictive power.

In conclusion, familiarity with a particular melody serves to raise the initial activation necessary to recognize it, whereas the time course of melody recognition relates to the occurrence of distinctive melodic characteristics. Whereas both pitch and timing information contribute to melody recognition, accumulated rhythm and momentary metric stress together account for most of the variance in POR. The combined contributions of accumulated and momentary measures suggest that we are not reliant on only one form of representation. Differences between the weight of these measures for nonmusicians and musicians might point to a shift from momentary information to a cumulative threshold with increased exposure to music. Future research could systematically track such a shift by using an artificial musical grammar to examine the importance of momentary and accumulated information in predicting the point of pattern recognition arising from different types of information from, and degrees of exposure to, the grammar. Such work has implications for our understanding of the representation of dynamic patterns in general and of the cognitive shifts associated with increased expertise.

AUTHOR NOTE

The present research was conducted largely during the author's post-doctoral research scholarship in the Cognitive and Systematic Musicology Laboratory at Ohio State University from 2003 to 2004. The work arose as a result of discussions with Professor David Huron, who helped to design the study, contributed many ideas, and provided all of the information measures. Thanks to Pia Gyldenkaerne for data processing, to David Pederson for statistical advice, and to anonymous reviewers who provided insightful comments on a previous draft of this article. Address correspondence to F. Bailes, MARCS Auditory Laboratories, Building 5, Bankstown Campus, University of Western Sydney, Locked Bag 1797, Penrith South DC NSW 1797, Australia (e-mail: f.bailes@uws.edu.au).

REFERENCES

- BARLOW, J., & MORGENSTERN, S. (1948). *A dictionary of musical themes*. New York: Crown.
- BHARUCHA, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, **5**, 1-30.
- BHARUCHA, J. J. (1999). Neural nets, temporal composites, and tonality. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 413-440). San Diego: Academic Press.
- CARLSEN, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology*, **1**, 12-29.
- CUDDY, L. L., COHEN, A. J., & MILLER, J. (1979). Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology*, **33**, 148-157. doi:10.1037/h0081713
- DALLA BELLA, S., PERETZ, I., & ARONOFF, N. (2003). Time course of melody recognition: A gating paradigm. *Perception & Psychophysics*, **65**, 1019-1028.
- DAVIES, J. B., & YELLAND, A. (1977). Effects of two training procedures on the production of melodic contour, in short-term memory for tonal sequences. *Psychology of Music*, **5**, 3-9. doi:10.1177/030573567752001
- DEMOREST, S. M., & KIM, D. A. (2004, August). *The role of pitch and rhythm cues in identifying mismatched melodies: Comparing experts and novices*. Paper presented at the 8th International Conference on Music Perception and Cognition, Evanston, IL.
- 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.
- HALPERN, A. R., MÜLLENSIEFEN, D., & WIGGINS, G. (2008, August). *Modelling memory responses in a melody recognition task*. Paper presented at the 10th annual International Conference on Music Perception and Cognition, Hokkaido University, Sapporo, Japan.
- HÉBERT, S., & PERETZ, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory & Cognition*, **25**, 518-533.
- HURON, D. (1994). *The Humdrum Toolkit: Reference manual*. Menlo Park, CA: Center for Computer Assisted Research in the Humanities.
- HURON, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
- HURON, D., & OLLEN, J. (2003). Agogic contrast in French and English themes: Further support for Patel and Daniele (2003). *Music Perception*, **21**, 267-271. doi:10.1525/mp.2003.21.2.267
- IDSON, W. L., & MASSARO, D. W. (1978). A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics*, **24**, 551-565.
- KORIAT, A., & LEVY-SADOT, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 34-53. doi:10.1037/0278-7393.27.1.34
- KRUMHANS, 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. doi:10.1037/0033-295X.89.4.334
- MCAULEY, J. D., STEVENS, C., & HUMPHREYS, M. S. (2004). Play it again: Did this melody occur more frequently or was it heard more recently? The role of stimulus familiarity in episodic recognition of music. *Acta Psychologica*, **116**, 93-108. doi:10.1016/j.actpsy.2004.02.001
- PEARCE, M. T., & WIGGINS, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception*, **23**, 377-405. doi:10.1525/mp.2006.23.5.377
- PEYNIRCIOĞLU, Z. F., TEKCAN, A. I., WAGNER, J. L., BAXTER, T. L., & SHAFFER, S. D. (1998). Name or hum that tune: Feeling of knowing for music. *Memory & Cognition*, **26**, 1131-1137.
- PFORDRESHER, P. Q. (2003). The role of melodic and rhythmic accents in musical structure. *Music Perception*, **20**, 431-464. doi:10.1525/mp.2003.20.4.431
- REDER, L. M., PAYNTER, C., DIANA, R. A., NGIAM, J., & DICKISON, D. (2008). Experience is a double-edged sword: A computational model of the encoding/retrieval trade-off with familiarity. In A. S. Benjamin & B. H. Ross (Eds.), *Skill and strategy in memory use* (pp. 271-312). San Diego: Academic Press.
- SCHAFFRATH, H. (1995). The Essen folksong collection [Online database]. Menlo Park, CA: Center for Computer Assisted Research in the Humanities.
- SCHELLENBERG, E. G., IVERSEN, P., & MCKINNON, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review*, **6**, 641-646.
- SCHMUCKLER, M. A. (1997). Expectancy effects in memory for melodies. *Canadian Journal of Experimental Psychology*, **51**, 292-305. doi:10.1037/1196-1961.51.4.292
- SCHULKIND, M. D., POSNER, R. J., & RUBIN, D. C. (2003). Musical features that facilitate melody identification: How do you know it's "your" song when they finally play it? *Music Perception*, **21**, 217-249. doi:10.1525/mp.2003.21.2.217
- STEVENS, C., & LATIMER, C. (1997). Music recognition: An illustrative application of a connectionist model. *Psychology of Music*, **25**, 161-185. doi:10.1177/0305735697252006
- TABACHNICK, B. G., & FIDELL, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.
- WHITE, B. W. (1960). Recognition of distorted melodies. *American Journal of Psychology*, **73**, 100-107.

NOTES

1. In the present article, the terms *recognizing* and *identifying* are used interchangeably with respect to melody. The former encompasses the latter, since it allows for the accurate retrieval of a melody from memory without necessarily being able to label or name it.
2. In the present study, melody recognition is understood to arise from familiarity with a particular melody and to include, but not require, the ability to name it or provide other identifying information in addition to knowing "how it goes."
3. Unfamiliar or relatively unknown melodies were included in order to maximize the variance in familiarity ratings, but it should be noted that, inevitably, many could not be recognized by participants in the subsequent melody recognition experiment.
4. Being primarily designed to maximize the variance of both types of measure (familiarity and distinctiveness), these classes were not initially tested for orthogonality. However, familiarity ratings from the melody recognition experiment are uncorrelated with measures of distinctiveness, and are thus orthogonal.
5. Although the experiment was run by the author, who was not naive to the hypotheses, the potential for a biased assessment of the POR was reduced by using stimulus notation during the experiment that was not marked by category (e.g., high familiarity or high distinctiveness), by ignorance as to the information values and, hence, distinctiveness associated with any particular note, and by the systematic cutoff of a trial after 10 notes.
6. Very few misidentifications occurred, even though many melodies begin with identical pitch and temporal patterns.

SUPPLEMENTAL MATERIALS

Supplemental interval and scale-degree-dyad data, as well as complete lists of the melodies chosen to study the familiarity and distinctiveness variables, may be downloaded from <http://mc.psychonomic-journals.org/content/supplemental>.

(Manuscript received June 18, 2009;
revision accepted for publication January 8, 2010.)