# 5

# Melody: Attention and Memory

*Music is a kind of kaleidoscope. . . . It brings forth a profusion of beautiful tints and forms, now sharply contrasted and now almost imperceptibly graduated; all logically connected with each other, yet all novel in their effect; forming . . . a complete and self-subsistent whole.*

(Hanslick, 1854/1957, p. 48)

*Just as the eye completes the lines of a drawing which the painter has knowingly left incomplete, just so the ear may be called upon to complete a chord and cooperate in its resolution, which has not actually been realized in the work.*

(Stravinsky, 1956, p. 36)

## INTRODUCTION

When we listen to music, our attention fluctuates, focusing first on one aspect and then another in the kaleidoscope. What we remember of a piece depends greatly on what we have attended to in listening. Sometimes our attention is "grabbed" by a salient feature in the stream of sound—a trumpet solo louder and more brilliant than the surrounding texture, for example. We can also direct our attention to features that are cognitively important even though they may not be salient in the stimulus, as when we follow the progress through a complex texture of an inner line that carries important musical information. In such cases, our attention is guided by knowledge structures developed in our experience of the world, called schemata (F. C. Bartlett, 1932; Neisser, 1976). A schema may embody general knowledge of stimulus properties common to many pieces of music, as in the case of our knowledge of tonal scales described in Chapter 4. A more specific type of schema may also embody knowledge of tonal relationships within a particular melody, like a melodic contour. Schemata guide our expectations of what will happen next, and, hence, what we attend to and remember. Schemata and the expectations they generate are always more general than the sounds that are actually heard. That is, expectations are rarely so specific and so dominant as to lead us into perceptual errors. Nevertheless, it is true that two people with two different sets of expectations can listen to the same stimulus and perceive different things.

An example will serve to clarify several of these points. Figure 5.1A shows the notes of the familiar "Frère Jacques" temporally interleaved with the notes of "Twinkle, Twinkle." The two melodies are in different pitch ranges. When the whole pattern is presented rapidly as in Example
5a (7.5 notes/sec), each melody can be heard clearly. The pitch difference 5a
between the melodies is a salient stimulus feature that we can use to focus our attention on either the upper line or the lower. You can listen at will to one melody or the other, but not to both at once. Virtually any dimension of the stimulus pattern could serve as a basis for focusing attention (Dowl-
ing, 1973a). In Example 5b (Figure 5.1B), the pitch ranges of the two 5b
melodies overlap, but the timbres of the two melodies have been changed (indicated in the figure by filled and open notes). You can still hear either
melody quite clearly. The same is true of Example 5c, in which the two 5c
melodies are distinguished only by loudness. Notice that you can attend to the softer melody almost as well as to the louder. Spatial separation can also differentiate the two melodies, as discussed in Chapter 2. In Example
5d, the melodies are fed through separate stereophonic channels, and it is 5d
easy to focus on one or the other.

Each of the foregoing examples provides the opportunity for selective attention on the basis of some stimulus dimension: pitch, timbre, loudness, or location. In music not constrained by the one-note-at-a-time restriction of these interleaved melodies, temporal dimensions such as rhythm and meter can be used to distinguish musical lines that the com-

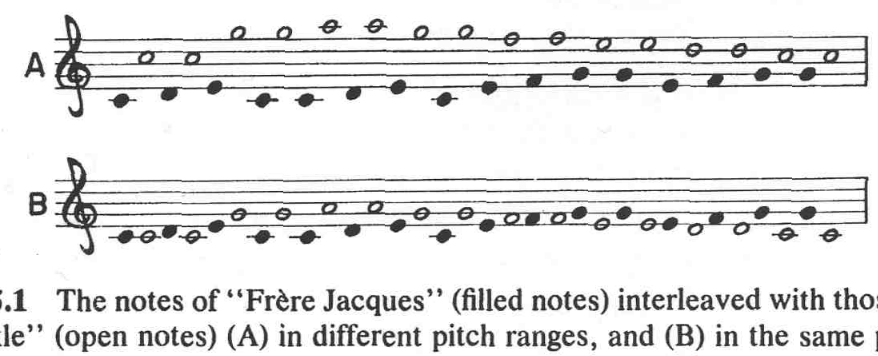


**Figure 5.1** The notes of "Frère Jacques" (filled notes) interleaved with those of "Twinkle, Twinkle" (open notes) (A) in different pitch ranges, and (B) in the same pitch range.


**Figure 5.2** Transcription of an excerpt from Mozart's finale to Act I of *Don Giovanni*, in which three groups of musicians play different dances simultaneously in different parts of the stage.

poser wants to clarify. Composers often make different lines proceed at different rates. Examples abound, including the running figure versus the hymn tune in Bach's chorale prelude "Jesu, Joy of Man's Desiring," as well as the rapid-fire staccato of Maddalena's griping versus the soaring sublimation of Gilda's resignation in Verdi's Quartet from *Rigoletto*. Differences of rhythm and rate of presentation are especially useful in polyphonic music, where the composer wants to give the listener a choice from among a variety of important musical lines, and in opera, where the composer may desire the simultaneous presentation of several points of view. Mozart, in the finale to Act I of *Don Giovanni,* provides a tour de force of temporal differentiation when he puts two bands on the stage in addition to the orchestra in the pit and has them all play different dances in different meters. Figure 5.2 conveys some of the complexity of Mozart's score. Spatial location, pace (the rate at which the notes go by), and metric organization all provide the listener with the means of focusing on one of the dances and ignoring the others.

Now let us return to the example of the interleaved melodies and demonstrate what happens when two familiar melodies are interleaved but without any simple stimulus feature to distinguish them. The pattern would be like that of Figure 5.1B, but with all the notes the same color.
5e Example 5e presents in just that way a new pair of melodies we think are familiar to you. We expect that the example sounds like a meaningless jumble of notes. Now listen again to "Frère Jacques" and "Twinkle,

Twinkle,'' but this time interleaved in the same pitch range without any physical difference in features (Example 5f). We trust you hear either melody, depending on which one you attend to. As the example is repeated, try shifting your attention back and forth from one melody to the other. You can hear the melodies clearly in Example 5f but not in Example 5e, yet the only difference between them is that you know which melodies to listen for in 5f. With 5f, you have a pattern of expectancies, a schema, that you can match against the stimulus to check for the presence of the expected melody. With 5e, you do not know which schemata to use and so cannot discern the tunes. Dowling (1973a) verified this phenomenon in an experiment by presenting listeners with pairs of interleaved melodies such as those in Examples 5e and 5f, preceding each pair with a true or a false label. (In 5f the true label is ''Twinkle, Twinkle'' and a false label might be ''On Top of Old Smoky.'' Is one of the tunes in 5e ''Happy Birthday''?) With true labels listeners almost always reported hearing the target tune. With false labels listeners almost never reported hearing the labeled tune even after 20 repetitions, nor did they correctly recognize either actual melody. We suppose that when listeners did recognize a mislabeled melody it was because they guessed correctly which schema to try matching to the stimulus. (We will tell you later what the tunes in 5e are.) **5f**

Neisser (1979; Bahrick, Walker, & Neisser, 1981) describes a visual analog of the interleaved-tunes phenomenon. If videotapes of two different games are superimposed on one screen, viewers find it easy to follow one series of events (e.g., a game of catch) and ignore another (e.g., handclapping). When viewers are asked to press a button each time the ball is thrown, for example, their accuracy is affected very little by the presence of the other game on the same screen. This is true even when the ignored game is another game of ball. As with the interleaved melodies, a schema of expectancies provides the person with the means of focusing attention on one series of events and ignoring other series of events even though the two are thoroughly intermingled in the stimulus display. As Bahrick et al. (1981) put it:

> Perception takes place when appropriate schemata are actively and continuously tuned to the temporally extended information that specifies an individual event. Irrelevant events present information, too, but remain unperceived simply because no such active tuning occurs with respect to it. (p. 378)

## The Nature of a Melody Schema

Both with interleaved tunes and with superimposed ball games, perceivers can follow the target events, provided they know which schema to use. With interleaved tunes, the critical events are so thoroughly embedded in the context that a rather specific schema (of a particular melody) is required to sort them out. (Incidentally, the tunes in Example 5e are "Mary Had a Little Lamb" and "Three Blind Mice.") This raises the issue of what such a melody schema is like. Several considerations lead us to suppose that it is not likely to be a literal mental copy of the melody. An exact copy would have to be translated—expanded, contracted, and shifted both in time and in pitch—to fit any actual instance of the melody that might be perceived. That is, a familiar melody can be presented at any arbitrarily selected pitch level and at any tempo (within broad limits), and we can still recognize it. Therefore, it seems likely that a melody schema should represent more general higher-order information than specific pitches at specific temporal intervals. Dowling (1978c) suggests that the pitch information in melodies might be stored in a schema consisting of the contour—the pattern of ups and downs—of the melody, plus an indication of where that contour should be hung on a tonal scale. Evidence is equivocal on this issue, but our present guess is that the schema of a familiar tune is somewhat more specific than a contour. If such a schema is more specific than contour, two possibilities occur. One possibility is that the abstract representation of pitches in a melody is the sequence of (logarithmic) pitch intervals. The second possibility is that pitches are stored as a sequence of abstracted chromas (i.e., do–re–mi labels in a movable-do system). We review the evidence, which we believe at this writing favors the second interpretation.

Some of this evidence is based on a type of stimulus that is in a way complementary to the interleaved melodies you have just been hearing. In those, perceptual confusion is produced by mixing two melodies in one pitch region. In our next group of stimuli, confusion is produced by scattering the pitches of one melody across several octaves while preserving
**5g** their chromas. This is illustrated in Figure 5.3A and Example 5g. The wide leaps of pitch make the melodic line hard to follow. Deutsch (1972) found that such octave-scrambled melodies are very difficult to recognize. Dowling (1978b) found that giving listeners true and false labels for such melodies produced results similar to those described above for interleaved melodies, though displaying lower accuracy. Correct labels produced 80% correct recognition, while incorrect labels misled listeners 25% of the time, giving an overall rate of 77% correct. Leaving the contour intact in the octave-scrambled version is some help to listeners if they are informed of the presence of the contour (Dowling & Hollombe, 1977; Idson & Massaro, 1978). Dowling and Hollombe found performance in that case of about 65% correct. We suppose that both the label and the contour can serve to retrieve a particular melodic schema from among the

(A)

(B)

**Figure 5.3** Melodies in which successive notes have been assigned to different octaves (A) preserving chroma (Deutsch, 1972), and (B) with chroma distorted.

many stored in long-term memory. Before positive recognition is reported, however, the schema is checked for chroma matches in the stimulus. If the chromas of the stimulus do not match those of the schema, the proposed schema is rejected. Kallman and Massaro (1979) found that when they preserved the contours of octave-scrambled melodies but altered the chromas by 1 or 2 semitones, correct recognition fell from about 65% (with both contour and chroma) to about 10% (with contour but not chroma). Example 5h presents a chroma-altered version of "Yankee Doodle," shown in Figure 5.3B. This suggests that chroma information is an integral part of the long-term memory representation (schema) for a familiar tune. The fact that recognition is much better than chance when chroma is preserved in these octave-scrambled melodies indicates that some abstraction of chroma, rather than interval sizes, is important, since intervals have been destroyed by the scrambling. And the fact that contour alone is not sufficient for recognition suggests that chromas are included in the schema and not just optionally accessible when needed (as, for example, via a separate scale schema as in Dowling's, 1978b, model). 5h

In the following, we present evidence supporting the notion that schematic representations of familiar tunes in long-term memory consist of (rhythmically organized) sets of relative pitch chromas and that such representations can be accessed by means of labels and such global melodic features as contour. We also present evidence that the higher-order tonal-scale schema *does* function independently of particular melodies, as Dowling (1978b) proposed. And we look at short-term (episodic) memory, long-term (semantic) memory, and cognitive development with respect to the roles played by pitch, intervals, contour, and tonal scales. Figure 5.4a shows the tune "Pop Goes the Weasel" (A) with various comparisons (B–E) that illustrate the effects of preserving or altering those features of the original. (Example 5i). In (B), the pitches have all been changed by trans- 5i

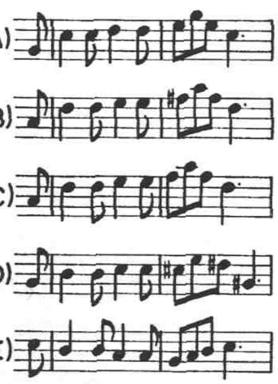


**Figure 5.4** (A) The beginning of "Pop Goes the Weasel" with various comparisons that (B) transpose the same intervals to a new key, (C) imitate the same contour with intervals changed, (D) depart from a tonal scale, and (E) alter the contour.

posing the tune to a different key, but the contour and intervals are the
5j same as in the original (Example 5j). In (C), some of the intervals have
been altered, and the tune no longer sounds exactly like "Pop Goes the
5k Weasel" (Example 5k). (C) is still within a tonal scale, but (D) uses
pitches outside any one tonal scale and is atonal while still preserving the
5l original contour (Example 5l). In (E) the contour is changed, and the tune
5m sounds very different (Example 5m). We begin our review of the evidence
by considering how pitch is perceived and remembered in a melodic context.

## MEMORY FOR MELODIC FEATURES

### Pitch

Though even novel, atonal melodies are easily recognized when repeated at the same pitch level (the A–B comparison in Figure 5.4), memory for single pitches is affected markedly by putting them into musical context. Krumhansl (1979, Experiment 3) found that pitches from a tonal scale were remembered well when followed by a context of pitches drawn from the same key, while a context of atonal pitches led to poorer memory for the target pitch. On each trial of the experiment, Krumhansl presented the listeners with a standard tone (for example, a G) lasting 0.5 sec. The standard was followed immediately by seven interference tones at a rate of two per second. The interference tones were either from the same tonal scale as the standard (e.g., C–E–A–F–D–B–C) or were an atonal sequence not in any key (e.g., C♯–E–A–F–D♯–B–C♯). Following the interference tones was a 1.5 sec pause and then a comparison tone that was either the same pitch as the standard or differed from it by 1 semitone. Listeners performed better than 95% correct when the interference tones were from the same key as the standard, while performance fell to about 80% correct with atonal interference. The opposite pattern occurred when the listener was trying to remember a standard tone *outside* the tonal scale of tonal interference tones (e.g., a G♯). In that case, tonal interference disrupted memory for the standard outside the key, driving performance below 80%. Atonal interference was not nearly so disruptive, leaving performance at about 90% correct.

What seems to be happening here is not that atonal contexts are disruptive per se, but rather that if the listener is trying to remember a standard pitch as a chroma in a particular key, the atonal context hurts performance. If the listener is trying to remember a pitch foreign to a key, then a context drawn from that key is disruptive and the atonal context is not. The tonal context appears to cause a shift in the listener's internal frame of reference when the interpolated tones are drawn from a different set than the one that incorporates the standard tone. This interpretation was explored further in work done by Kirk Blackburn in Dowling's laboratory. Blackburn used two types of tonal context rather than tonal and atonal and made the listener's task more difficult than Krumhansl's by asking listeners to *imagine* the standard tone. In other respects, the procedure was very similar to Krumhansl's. To aid the listener's imagination, Blackburn played part of a major scale leading up or down to the tonic (e.g., G–A–B or F–E–D), leaving it to the listener to imagine its completion (in this case, C). Five interference tones followed, either from the same key as the target (e.g., G–A–D–E–F) or from a distant key (B or F♯ major, e.g., F♯–G♯–A♯–D♯–F♯). The comparison tone was either the imagined target (C) or a semitone removed from it (B or C♯). When the interpolated tones were in the same key as the imagined tonic and its scale, performance was around 75% correct. When the interpolated tones suggested a different key, performance was worse than chance—around 40%.

The interpretation that distant-key interference caused a shift in the listener's schematic frame of reference is supported by the pattern of errors when the comparison tone differed from the imagined tonic C. The B could have come from either C major or the distant key (B major or F♯ major). The C♯, however, could have come only from the distant key. False-positive recognitions of the B were about equal for the two types of interference, while false positives for the C♯ were primarily the result of distant-key interference. The C♯ sounded very natural when it followed a series of tones with which it could combine in a major scale, whereas it sounded strange and was easy to reject when it followed the C-major scale to which it was foreign.

Memory for a pitch can be altered by contextual shifts other than those involved in the tonal structure. Pitch shifts of notes in brief atonal melodies affect memory for the pitch of neighboring notes. Guilford and Hilton (1933) used pairs of melodies from Seashore's (1919) test of melodic memory. The pitch of one note of the melody was changed upon repetition. Listeners reported hearing changes not only in the actually altered tones, but also shifts (in the same direction) of neighboring tones that had not been altered. In a second study, Guilford and Nelson (1936) repeated melodies without altering any pitches, and listeners still reported hearing pitch shifts. (The atonality of the melodies probably contributed to the difficulty of accurate pitch judgment.) Guilford and Nelson (1937) simplified the task by using three-note melodies containing pairs of identical or adjacent pitches, plus another note separated in pitch from the pair. The three notes could occur in any order, and the listener's task was to say whether the second note of the similar pair was the same as, or higher or lower than, the first. Guilford and Nelson found that the second note of the pair tended to be shifted away from the note that was different in pitch. For example, in the sequence C♯–G–C♯ (with the G higher) the second C♯ was judged lower in pitch than the first. It is as though the listener's internal standard for the pitch C♯ had been shifted upward by the occurrence of the G, and the second (actual) C♯ judged flat by comparison.

Dewar, Cuddy, and Mewhort (1977) provide further evidence of the importance of a tonal scale schema in memory for pitch. They presented listeners with seven-note sequences that were either tonal or atonal. Then they presented a pair of tones, one of which had occurred in the original sequence, and asked the listeners to tell which tone they had heard before. Performance was better with tonal sequences than with atonal (81% vs. 77%). Dewar et al. also included a condition in which the comparison stimuli consisted of the whole seven-note sequences, either intact or with one note changed. Listeners found this task much easier, achieving 99% correct with tonal sequences and 91% with atonal. We can conclude two things from this: (1) The additional information in the whole sequence was useful in judging the accuracy of the single pitch, and (2) this information was especially useful with tonal sequences.

In an extension of this line of work Cuddy, Cohen, and Miller (1979) tested memory for tonal three-note fragments. Listeners were supposed to notice a change in one note of a fragment when it was presented and tested in isolation, or with the addition of a context of two preceding and two following notes. The altered note either remained within the tonal scale of the other notes or departed from it. In comparison to detection of note changes in the fragment alone, addition of a strongly tonal context led to significantly better detection of alterations that departed from the tonal scale. Addition of an atonal context led to worse performance in the detection of alterations whether within the tonal scale or not. The effects of context depend on the degree to which context invokes the listener's scale schemata, a point to which we return below.

In summary, we have seen that the context in which a pitch is heard affects memory for that pitch, and in particular, that tonal scale context can aid in memory for context-compatible pitches and aid in the detection of incompatible ones. Contexts that include pitches outside the scale schema of an inferred tonal scale can interfere with accurate memory and cause systematic errors of judgment. A practical application of this principle is found in choral singing. When a section of a chorus has several measures rest, reentry on the correct pitch is often facilitated by the reentry pitch's being the same as the last pitch sung. As long as the piece stays in the same key during the rest, choristers remember the pitch well. However, if the piece modulates to a new key the entrance is more difficult, even though the pitch is the same.

In the broader scheme of remembering music, remembering a melody involves more than just remembering a series of unrelated pitches. Melodies have global features that pertain to the whole pattern, and one of these is contour. We now turn to a discussion of memory for melodic contour.

## Contour

Contour refers to the pattern of ups and downs of pitch from note to note in a melody. The importance of contour in recognition is disclosed in experiments in which comparison melodies sharing the same contour as the original (such as those in Figure 5.4C, D, and E) are easily distinguished from those that do not (as in Figure 5.4F). The relative importance of contour in comparison to other features of melodies is shown by the degree to which the (C–D–E) versus (F) discrimination is strong and the discriminations among melodies like (C), (D), and (E) (that differ among themselves in other features such as pitch intervals and tonality) are weak. Contour is an especially important feature of melodies in immediate recognition where the exact relationship between a melody and the scale schema has not been thoroughly established, as well as with atonal melodies in which there is no scale schema to relate the melody to.

The dominance of contour in the immediate recognition of atonal melodies is illustrated in a study by Dowling and Fujitani (1971, Experiment 1). They presented listeners with pairs of five-note atonal melodies like those shown in Figure 5.5. The comparison melody was either an exact transpo-

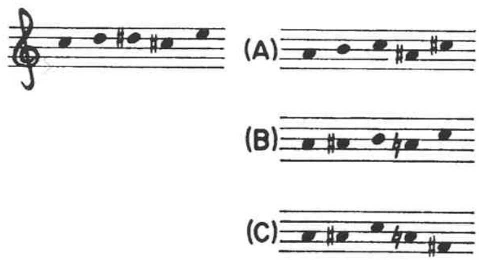


**Figure 5.5** Atonal melodies like those used by Dowling and Fujitani (1971, Experiment 1): (A) exact transposition, (B) same-contour imitation, and (C) different-contour comparison.

sition of the original to a new pitch level (Figure 5.5A), an imitation of the original that preserved contour but not interval sizes (Figure 5.5B), or a comparison with a different contour (Figure 5.5C). Listeners found it relatively easy to distinguish between either transpositions (A) or contour-preserving imitations (B) and the different-contour melodies (C), achieving between 85% and 90% correct. Listeners found it almost impossible to distinguish between transpositions (A) and same-contour imitations (B), however, performing at around the chance level of 50%. It appears that these listeners, who had at most only moderate amounts of musical training, based their judgments almost entirely on contour similarity and were unable to detect changes in interval sizes in these atonal melodies.

In a similar experiment Francès (1958, Experiment 9) asked listeners to distinguish transpositions of brief melodies from same-contour imitations. Listeners found the task much harder with atonal than with tonal melodies, suggesting that they did not succeed in remembering the intervals between the notes in the atonal melodies and were confused by imitations that had similar intervals. In that case, contour was the dominant feature. Francès' study had the virtue of using more natural sounding melodies than most studies in this area—melodies having interesting rhythmic patterns. Though Francès' study had the limitation of using only four different melodies over and over again, the fact that its results converge closely with other findings leads us to have confidence in them. We return below to the role of tonal scale schemata in melodic memory. For the present, it seems clear that contour is an especially important feature in the recognition of atonal melodies.

Contour is also important in the immediate recognition of novel melodies in cases where the tonal scale framework remains constant. Dowling (1978b) replicated Dowling and Fujitani's (1971) experiment, but this time used tonal melodies. Figure 5.6 illustrates the types of comparison melodies in this study. Note that the tonal imitation (C) remains in the tonal key of the original; that is, it is constructed from notes of the same tonal

**Table 5.1**
Area under the Memory Operating Characteristic, as Estimated Percentage Correct

| Group | Transposition compared to: Tonal imitation | Atonal imitation | Different contour |
|---|---|---|---|
| Dowling (1978b) | | | |
| Inexperienced | 49 | 59 | 81 |
| Experienced | 48 | 79 | 84 |
| Dowling and Fujitani (1971) | — | 53 | 89 |

scale. Comparison (C) preserves the same *diatonic* intervals (measured along the tonal scale) as (A) while altering the *chromatic* intervals. Listeners were unable to distinguish between (A)–(B) pairs and (A)–(C) pairs in this experiment, showing that contour is the dominant feature in those comparisons. Example 5n presents first an (A)–(B) pair and then an (A)–(C) pair so that you can hear how similar the (A)–(C) melodies sound. The data are shown in Table 5.1 along with the corresponding condition of Dowling and Fujitani's experiment. Discrimination between transpositions and same-contour tonal imitations is around the chance level of 50%, while discrimination between transpositions and different-contour melodies is better than 80%. Notice that while Dowling and Fujitani's listeners could not distinguish transpositions from imitations where both were atonal, Dowling's (1978b) listeners could do so at better than chance levels where the transposition is tonal and the imitation atonal. Tonality itself can be used as a cue, and naturally enough, the more experienced listeners were better at using it (79% vs. 59%). We return to this study in our discussion of tonality, but for present purposes, it is clear that listen- **5n**

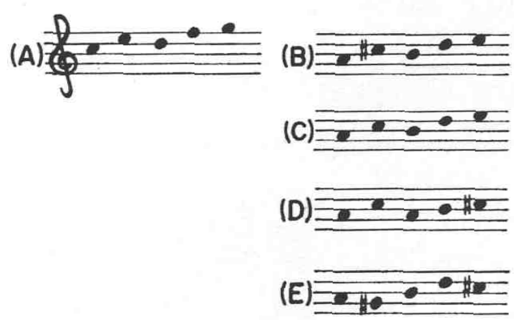


**Figure 5.6** Tonal melodies used by Dowling (1978b): (A) initial melody of trial, (B) exact transposition, (C) tonal imitation with different intervals, (D) atonal imitation with different intervals and pitches outside tonal scale, and (E) different-contour melody.

ers have difficulty discriminating tonal imitations in the same key from transpositions and that this leaves contour the dominant melodic feature in determining the listeners' responses.

Contour is an important feature in the recognition of familiar melodies. White (1960) and Dowling and Fujitani (1971, Experiment 2) demonstrated that listeners can recognize distorted versions of familiar tunes in which the pitch intervals between notes are changed while the contours are preserved. Dowling and Fujitani used a set of tunes of which the first two phrases could be regularized into the same rhythm, thus eliminating rhythmic pattern as a cue. Undistorted versions of these tunes were recognized almost perfectly, while distortions in which contour had been destroyed were recognized only 30% of the time (a little better than chance). When the distortion preserved contour information, performance rose to about 60% correct. Performance improved somewhat more if *relative* interval size information was included with the contour, that is, if larger intervals remained larger after distortion and smaller ones remained smaller.

## Key Distance

The relative importance of contour information in melody recognition varies with tonal scale context. The ease with which listeners can distinguish between transpositions of a melody and same-contour different-interval imitations depends upon the relationship between the key of the test melody and the key of the original. In Chapter 4, we introduced the notion of distance between keys and reviewed evidence from mulitdimensional scaling that key-distance has psychological reality. Here we discuss the effects of key-distance on melody recognition. Distance between keys is measured in music theory by the number of different pitches in the tonal scales of the two keys. This can range from one out of seven scale pitches at the near end to six out of seven at the far end. This is illustrated in Figure 5.7 in which three major scales are shown as selections of pitches from the chromatic scale of 12 semitones in the octave. The C-major and the D-major scales are relatively close, sharing all but two of their pitches, while the C-major and B-major scales are distant, having only two pitches in common.

In a series of experiments, J. C. Bartlett and Dowling (1980) manipulated the key relationships between the initial melody and the comparison melody in a pair, using an immediate recognition paradigm very similar to those described above. Comparison melodies were of the types shown in Figure 5.6 but in a variety of keys. For example, on a *tonal imitation* trial, the first melody of a pair might be in C major starting on the first degree of

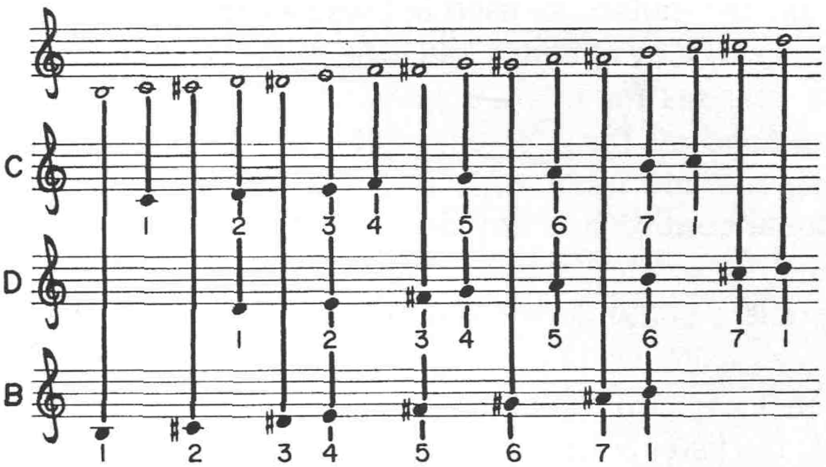


**Figure 5.7** Closely and distantly related keys seen as selections of pitches from the tonal material. Closely related keys (e.g., C and D) share more pitches than distantly related keys (e.g., C and B).

the scale, and the second melody might be in D major and shifted to begin on the third degree of the scale (F♯). That would be a *near-key tonal imitation.* A *far-key tonal imitation* might be an imitation in B major, starting on the third degree of the scale (D♯). In this sense, the tonal imitations used by Dowling (1978b), illustrated in Figure 5.6C, are *same-key tonal imitations.* Bartlett and Dowling (1980, Experiment 1) replicated Dowling's (1978b) study and obtained very similar results, with performance on recognizing transpositions and rejecting same-key imitations, atonal imitations, and different contour stimuli all falling within five percentage points of the results shown in Table 5.1. The additional result that Bartlett and Dowling found was that as key distance of imitations was increased from same to near to far, listeners were less confused by them and found them easier to reject. In subsequent experiments Bartlett and Dowling (1980) found that this key-distance effect was mainly due to listeners' better rejection of far-key imitations, rather than to better recognition of far-key transpositions.

This result suggests that listeners use schematic scale information in solving certain aspects of the transposition-recognition task. Since the exact pitch intervals between the notes of a novel melody are difficult to remember, the listener uses melodic contour in conjunction with the chroma set of the scale. Where the key of the comparison melody is very similar to the key of the original, imitations are hard to reject because they share the contour and chroma set of the original. When comparison melodies are shifted to a far key, the chroma set is different and no longer misleading. The interval information available in the listener's memory, though meager, is sufficient to reject imitations with greater than chance accuracy. (Note that it takes only one mismatched interval to reject an imitation, and the imitations used in these experiments each had several interval changes.) Key distance had little effect on recognition of transpositions. The reasons for this are probably complex. We consider below possible explanations for different effects of key distance on recognition of transpositions and imitations. But first we describe some additional effects of tonal context and key distance.

In the study described above, Cuddy, Cohen, and Miller (1979) tested whether listeners could detect alterations in three-note tonal sequences. The sequences were either presented alone or embedded in contexts that varied in tonal strength from atonal to strongly tonal. On each trial of the experiment, the listener heard a standard melody followed by a transposition *and* a transposition with one of its pitches altered. Each trial was presented five times, and each time the order of the two transpositions was randomized. The listener's task was to say which comparison melody was the accurate transposition. The transpositions were either to near or to far keys, and the altered notes either remained inside the new key or departed from it. As we would expect, the altered notes that went outside the key were especially easy to notice when there was a strong tonal context. What is at first sight surprising in the results of Cuddy et al. is a key-distance effect running in the opposite direction from that obtained by Bartlett and Dowling (1980). With alterations remaining inside the key of transposition, listeners were better at distinguishing between exact and altered transpositions in near keys than in far keys. (This was true with strong tonal context and without context, but not with atonal or weak tonal context.) We believe the difference in results between these two studies to be attributable to a difference in method. The method that Bartlett and Dowling used presents the original melody and the comparison just once, and it is likely that the listener is not able to shift effectively to the schema of the comparison. In that case, the listener is confused by near-key imitations because of failure to shift to a new key. Far-key imitations are not so confusing, because obvious violations prevent their interpretation in the original key. In contrast, the method of Cuddy et al., with its repetition of each trial, provides ample opportunity for the listener to shift to the key of the comparison. This is more effectively accomplished to near keys than to far, and so the listener performs better with the near keys. Both studies illustrate the importance of the tonal scale schema and key distance, but in different ways.

## Intervals and Chromas

Though contour is useful in the recognition of familiar tunes stored in long-term memory, it is clear that intervals (patterns of chromas) are much more important there than in the immediate recognition of novel melodies in the studies discussed above. Recall that in the study by Attneave and Olson (1971) discussed in Chapter 4, even nonmusicians could recreate the interval pattern of the NBC chimes quite precisely at arbitrarily chosen pitch levels. Accuracy in noticing distortions of intervals of familiar tunes is a common finding. Bartlett and Dowling (1980, Experiment 2), in their series of experiments on key-distance, tested immediate recognition of melodic phrases drawn from either familiar folk songs or unfamiliar pseudo–folk songs. In the difficult task of distinguishing transpositions from same-contour imitations, performance was much better with familiar than with unfamiliar melodies. The imitations of unfamiliar melodies were not so confusing as those used by Dowling (1978c), probably because they were rhythmically more interesting. Performance with unfamiliar melodies was around 70% (with chance at 50%). But with familiar melodies, performance leapt to about 90%, indicating very good recognition of intervals. This provides documentation of the ease with which listeners are able to reject interval-distorted versions of familiar tunes such as those you heard in Example 5k.

The recognition of a familiar tune is an example of the use of long-term (semantic) memory. Psychologists contrast long-term and short-term (episodic) memory (Lindsay & Norman, 1977). The use of short-term memory is illustrated by the recognition of a novel melody first presented immediately before the comparison. Generally speaking, short-term memory is thought to be limited in capacity to about seven items at a time and to hold information for periods of up to about 30 sec. That information needs to be written into long-term memory if it is to be remembered for a longer period. Long-term memory is viewed as virtually unlimited in capacity and as able to store items for indefinite periods of time. The difficulties with recognition of items in long-term memory arise largely from the problem of retrieval, that is, of finding the relevant memory record from among the immense number of records stored there. The question of *which* item one is searching for looms very large. In short-term memory, the question of which item one wants is usually not crucial. The items are already available to be tested for a match. These differences in memory processes lead to differences in the importance of the various features of a melody to be recognized, depending on whether long-or short-term is involved. For example, it seems likely that relatively specific information (such as chromas or interval sizes) is more important in long-term memory processes, where a melody has to be differentiated from a large number of similar alternatives, than in short-term memory where the few alternatives are already available. This is what Dowling and Bartlett (1981) found.

This result—the importance of chroma or interval information in long-term memory—at first surprised Dowling and Bartlett. They were sure of the importance of contour information in short-term memory and expected that contour would also dominate in long-term memory. They were also thinking of the way composers use contour similarity to give unity to a piece. The first movement of Beethoven's Fifth Symphony is a prime example of this type of writing. Beethoven's familiar four-note theme is first presented with an interval of 4 semitones between the last two notes and then repeated immediately with an interval of 3 semitones. During the next 30 sec, the theme is repeated over and over with intervals of 1, 3, 4, 5, and 7 semitones. The same contour recurs through all of these repetitions, and it seems unlikely that any listener could escape having it firmly engrossed on his or her memory by the end of the 8-min movement. Beethoven was relying on the listener's memory for contour to provide structure for his piece. Since those memories need to last over periods of minutes within the piece, and, in fact, last for years afterwards, they would seem to involve long-term memory processes.

Dowling and Bartlett (1981) set out to explore these possibilities with an experiment requiring memory for excerpts from Beethoven's String Quartets. They collected pairs of presentations of themes related in the same way as the first two presentations of the first theme in Beethoven's Fifth, that is, pairs in which the second presentation imitated the first with the same contour but different intervals and chromas. Listeners heard a series of 18 excerpts including one member of each pair. Then after a 5-min pause, the listeners were tested with a series that included exact repetitions of what they had heard before, imitations (the other members of the pairs), and completely different excerpts not heard before. Listeners were told to try to recognize the imitations, giving positive responses both to them and to the exact repetitions. Listeners succeeded at recognizing exact repetitions, distinguishing them from completely different excerpts with 75% accuracy. But listeners were unable to distinguish the same-contour imitations from different items, performing at chance level (50%). This was surprising in view of the great similarity of repetitions and imitations, which not only shared the same contour but were almost always drawn from consecutive passages in the same piece (as with the Beethoven's Fifth example) and thus had similar tempo, instrumental color, and loudness. Listeners failed to recognize imitations even when instructed to do so and when the relationship of imitations and repetitions was explained and illustrated. It seems that these moderately experienced listeners could use their long-term memories in recognizing the pattern of interval sizes (or chromas) in novel pieces of music but could not recognize contours. For contour to be effective in retrieval from long-term

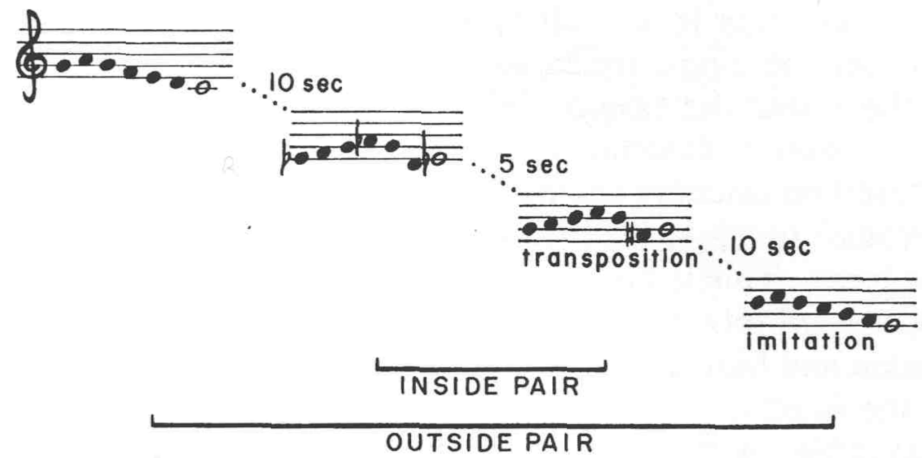


**Figure 5.8** Structure of a trial in Dowling and Bartlett's (1981) inside–outside procedure. Listeners responded following the third and fourth stimuli in each trial.

memory, it seems that the melody must be very well learned and familiar (Dowling & Fujitani, 1971, Experiment 2). It seems likely that a crucial feature of Beethoven's Fifth is that the theme is presented several times (not just once) within a span of time during which earlier presentations are still in short-term memory.

Dowling and Bartlett (1981, Experiment 4) explored the phenomenon of accuracy of interval information in long-term (vs. short-term) memory using a method that directly contrasted the two processes. Figure 5.8 shows the structure of a trial in this experiment. On each trial, the listener hears two pairs of melodies: an outside pair and an inside pair. The first member of the outside pair must be held in long-term memory while the listener evaluates the inside pair using short-term memory. Comparison melodies in both pairs included transpositions, tonal imitations, and different-contour items. Listeners were instructed to try to distinguish between transpositions and imitations, as well as between transpositions and different items. With the inside pairs, listeners distinguished transpositions and imitations from different items (75% and 72%, respectively) but not from each other. This is essentially the same pattern as found in the earlier short-term memory results shown in Table 5.1. With the outside pairs, performance was, of course, generally worse. But the relative difference between transposition and imitation recognition tended to widen. Recognition of transpositions was at 65%, but recognition of imitations was at 57% (as compared to different items). That is, recognition of transpositions fell 10% going from short-term to long-term memory, while recognition of imitations fell 15%. This suggests that the accuracy of interval or chroma information that listeners have for familiar melodies begins to develop during the first few times those melodies are entered into long-term memory. As Dowling and Bartlett (1981, p. 30) say, "While interval information is difficult to encode, it is apparently retained with high efficiency in long-term memory."

We believe that the ease of recognizing undistorted familiar tunes, as well as the relative importance of interval information in long-term memory, is based on listeners having stored in memory a sequence of relative pitch chromas (pitch levels in a tonal scale) rather than a set of intervals between tones. This is because when the intervals are distorted but the chromas are left intact, recognition is still possible, as in the experiment by Kallman and Massaro (1979), discussed above. When the chromas as well as the intervals are destroyed, then the melody becomes virtually unrecognizable. With unfamiliar tonal melodies, the chroma set of the scale becomes more important than the particular chromas of the melody itself, and confusions arise when the melody shifts in pitch but the scale does not. As long as the comparison melody preserves the contour of the original and uses the chromas of the same scale, the two will be confused. However, the degree to which chromas are represented in the listeners memory may depend upon individual differences in training in the use of tonal scale systems, as the following experiment suggests.

In this experiment, Dowling (1982a) transposed the interval patterns of melodies, in some cases leaving the pattern of chromas intact and in other cases changing it. Chroma is based on the place of a pitch in a tonal scale and so can be manipulated by changing tonal context. A melody using the first, second, and third degrees of the scale (do, re, mi) can have the same intervals as one using the fifth, sixth, and seventh degrees (sol, la, ti), though its chromas will be different. As long as a melody avoids the seventh and fourth degrees of the scale, it can be shifted from the do–re–mi position to the sol–la–ti position without any distortion of intervals. The shift of position can be determined by a chordal context pausing on the tonic and the dominant chord, respectively. (The tonic chord is based on the first scale position, do, and the dominant chord is based on the fifth scale position, sol.) This is illustrated in Figure 5.9. The original melody (in black notes in line A) begins and ends on C (do) in the key of C major, and that assignment of chroma values is established by the chordal context that precedes it. The comparison melodies are both transpositions of the same interval pattern to start and end on D. In (B) the chordal context establishes that D as first degree (tonic) of D major, and so (B) retains all but one of the same chromas as (A). In (C) the context makes D the fifth degree (dominant) of G major, and so the chromas are changed from do–re–mi (1–2–3) to sol–la–ti (4–5–6). The stimuli in Figure 5.9 can be heard
5o in Example 5o. This experiment was conducted as a continuous running memory task, in which the listener heard a succession of stimuli like those in Figure 5.9, and responded to each according to whether or not it had


**Figure 5.9** Examples of stimuli in which chroma is varied while interval pattern remains constant: (A) an initial melody of a pair; (B) a near-transposition with most of the same chromas and intervals, with one chroma changed; (C) an exact transposition with different chromas and same intervals (Dowling, 1982a).

occurred before in the series. Some of the stimuli were first members of pairs and had not been heard before, while others were second members of pairs that were transpositions or imitations of earlier items. In this case, imitations had the same contours as originals but had one pitch changed as in Figure 5.9B. Listeners were asked to respond positively only to exact transpositions. Since the lag between the presentation of an item and its mate was a little more than a minute and filled with responses to other items, this procedure presumably tapped long-term memory processes. Thus, it was not surprising from the evidence reviewed above that performance in distinguishing transpositions from imitations was relatively good. Musically inexperienced listeners performed in the 60–65% range on that task, both with chromas the same and chromas altered on test. However, moderately experienced subjects succeeded in that distinction only when chromas remained the same, performing around 65% correct. When chromas in the test stimulus were changed, experienced listeners fell to chance (50%) in distinguishing transpositions from imitations. (Experienced listeners were better overall at contour recognition—distinguishing transpositions *and* imitations from different stimuli—but worse overall at interval recognition.) This leads us to conclude that both inexperienced and experienced listeners store accurate interval information in long-term memory upon first hearing a novel melody. Inexperienced listeners store this information simply *as intervals,* perhaps using the type of automatic interval processing system suggested by Deutsch (1969, 1982). More experienced listeners make use of the tonal scale schema they have learned to store interval patterns as chroma patterns. This strategy works well so long as chroma remains constant when memory for intervals is tested. It breaks down when chroma changes. With recognition memory, we encounter the same types of individual differences in the use of scale schemata found by Krumhansl and Shepard (1979) with multidimensional scaling methods (discussed in Chapter 4). According to this explanation of Dowling's (1982a) results, inexperienced listeners should not perform appreciably worse than experienced ones in discriminating *atonal* transpositions from imitations, and, in fact, Dowling and Fujitani (1971, Experiment 1) found a correlation of only .23 between years of musical training and such performance.

From the above review, we can see that the features of contour, interval size (or chroma), and tonal scale system all play a role in the adult's perception and memory for melodies. Next we turn to the child's development of auditory cognition, and there, too, we find the same features important. But before leaving adult cognition, there are two points we wish to reinforce. (1) Different features of melodies can have different importance depending on task demands. Thus, contour is an important feature in short-term memory tasks, but not so in long-term memory, where many stored melodies share the same contour and interval sizes become important in differentiating among them. (2) Individual differences among people with different developmental backgrounds are important to consider in constructing a theoretical model of music cognition. This is especially true of behaviors that involve the use of elaborate schemata (like the tonal scale) that depend upon training.

## DEVELOPMENT

The same kinds of features that are important in adult information processing are also important in tracing the development of melody perception and memory. Different features become important in the child's behavioral repertoire at different ages, and this lends further support to the assertion that the various sets of features are psychologically distinct. In the remaining section of this chapter, we review evidence that melodic contour is already distinguishable early in infancy and that young infants can match the pitches of single tones. The child's ability to reproduce more than a phrase or two of a melody does not typically develop until sometime after the age of 2 years, and even then, the pitch of an extended melody generally wanders. Around the age of 5 years, the child becomes aware of changes of key and tonal center, and the melodies sung at that age wander less in pitch. It is not until a few years later, however, that the child is able to notice small changes of intervals or chromas in familiar melodies, a task at which the adult is very adept.

It is useful to think of the features involved in the processing of melodies as organized into the sorts of schemata we have been talking about, in fact, perceptuomotor schemata in Piaget's sense (Piaget & Inhelder, 1969). These schemata govern perception, memory, and production in both the child and the adult. We discuss numerous examples of the operation of these schemata in adult behavior in Chapter 4 as well as in this chapter. Tonal scale schemata are among the last to form in development; among the earliest are schemata for melodic contours. As we will see, young children deal with melodies in terms of contour. They notice changes in the contours of heard melodies, and when they sing, it is contour they control.

## Infancy

When an infant observes a new bloom or buzz in William James' (1890) blooming, buzzing confusion, it becomes startled. A number of researchers have used this startle reaction to distinguish the types of stimulus change that babies notice from those they do not. Specifically, the researchers observe the baby's heart rate. A sudden slowing of heart rate indicates startle. In a typical study, an auditory pattern is repeated over and over. On the first presentation, the baby is startled and exhibits heart-rate deceleration. Then as the baby gets used to the pattern, heart rate returns to normal—it adapts. After the baby adapts to the first pattern, the pattern is changed to a new one. The question is, What types of change in the pattern cause a new startle? The use of this paradigm requires babies who have sufficiently strong shifts in heart rate to indicate both startle and adaptation and who do not become too fussy or distracted in the experimental situation. Usually about half the 5-month-olds tested provide usable data.

Chang and Trehub (1977a) got 5-month-old infants to adapt to six-note melodies. The melodies had a tempo of 2.5 notes/sec and were atonal with rather large intervals between the notes. The first melody was repeated 30 times over a 5.5 min period while the infant adapted. Then Chang and Trehub shifted to a new pattern. The new melody was either a transposition of the first melody (up or down 3 semitones) or a permutation of the order of the notes of the transposition. The first of these new melodies preserved the contour of the original, while the second did not. In both cases, the shift in overall pitch level was the same, and new melodies contained exactly the same pitches. Chang and Trehub found that the babies reacted to the melody with the changed contour but not to the transposition. (The principal features of these results have since been replicated with more sophisticated methods by Trehub, Bull, & Thorpe, 1984.) It makes sense that contour should be a salient feature of tonal patterns for young infants just embarking on learning language. Nine-month-olds have been observed to babble using the sentence intonation contours of adult English. And in tone languages (spoken mostly in Africa and Southeast Asia), the pitch contour of a syllable is phonemic; that is, it makes a difference in meaning.

Though an infant of 6 months cannot sing a melody after an adult model, babies of that age can match the pitches of single notes quite accurately. Kessen, Levine, and Wendrich (1979) conducted a study in which an adult got the infant's attention and then sang a pitch (a note of the D–F–A triad well within the infant's vocal range). To the researchers' (initial) surprise, the infant attempted to sing the pitch back. After some practice at this task, which the infants apparently enjoyed, the babies became quite accurate. They matched the correct note of the triad two-thirds of the time, and when they produced a successful match, it was usually quite well in tune. There is also evidence that 3-month-old infants notice changes in pitch chroma, in the sense of responding similarly to pitches an octave apart. Demany and Armand (1984) used a method much like Chang and Trehub's. They found that after habituation to a brief sine-wave melody, infants were not startled by substitution of pitches an octave away from corresponding notes in the original but were startled by changes to pitches somewhat more or less than an octave away.

During the first 2 years of life, the child produces a number of behaviors that later become integrated into an overall pattern of musical behavior. Matching pitches, (noted above), recognizing specific tunes, singing single phrases from tunes, inventing phrases and spontaneous songs, and beating regular rhythmic patterns usually appear during this time. Our own experience as well as various reviews of observed onset times for these behaviors (Ostwald, 1973; Révész, 1954; Shuter-Dyson & Gabriel, 1981) convinces us that the ages at which children first do these things vary widely. Continued experience performing the various behaviors is perhaps more important to the child's development than age of onset. Regarding pitch matching, Kessen (1981) notes that it tends to disappear in subsequent development unless it continues to be practiced. We agree with Kessen that it is often helpful to the child's musical development to have a somewhat older sibling modeling these behaviors and providing stylistically accessible music in his or her own practicing. The sophisticated adult music the parents listen to or perform is generally far beyond the infant's comprehension. An older sibling provides a steady input of oft-repeated simple songs. As an illustration, one of us (WJD) has two daughters 2 years apart. The elder first labeled a tune (in the sense of consistently making a specific response to a specific tune) at 18 months. When the younger was 9 months old, the older child developed a passion for "Old Macdonald Had a Farm," with its repeated chorus "ee-ai-ee-ai-oh." She sang it alone or with others several times a day for a month, and then the frequency tapered off to several times a week. At around 11 months, the younger daughter began to respond with some variant of "ee-ai-ee-ai-oh" whenever she heard the first phrase of "Old Macdonald." Tested systematically at 12 months, she rarely generalized to the openings of other familiar tunes. Intensive exposure to a tune she could comprehend seems to be a major factor in her use of tune labeling 6 months earlier than her sister. It is also important that the experience with the tune occurred in a socially meaningful setting. That is, when she heard the tune, it was from members of her family, usually inviting her participation by glances and nods. Her family was obviously having a good time with the tune, and so she approached it as a fun game to learn, just as the babies did in the Kessen et al. (1979) pitch-matching study.

## Childhood

As children progress through the second year of life, they sing more and more coherently. Infants in their first year typically do a good deal of vocal play, exploring the pitch and dynamic ranges of their voices and the various timbres that are possible. During the second year, they do not leave off vocal play, but they also begin to produce patterns that an adult would recognize as coherently organized songs. These songs consist of short phrases, often just one phrase repeated over and over at different pitch levels. The pitch wanders without regard for any stable key but almost always moves by discrete steps from one focal pitch to the next. Some of the pitch intervals seem to follow adult models, but others do not. In the songs we have observed (Dowling, 1982b, 1984), no one interval seems to predominate. Some observers have been impressed with a tendency to descending minor thirds (3 semitones; e.g., Moog, 1976), even trying to see a cross-cultural universal there. In our opinion, the evidence is weak. If a mean of all descending interval sizes were taken from a large sample of 2-year-olds' songs, we would not be surprised if that mean were somewhere close to a minor third. The variance in the samples we have observed, however, is large enough to argue strongly against children of that age having a stable interval of a minor third or any other size.

Example 5p was produced by a girl just turning 2 years. It is typical of **5p**
her songs between the ages of 18 and 30 months. It is elaborate relative to most of her songs, having two phrase contours that alternate, rather than just one repeated contour. The words, her own invention, repeat. (Not all her songs had definite words—some went "la, la, la," and others had

vocalizations of the "ee-ai-ee-ai-oh" sort.) This song is a member of a family of spontaneous songs produced around this age. The simplest member of the family consisted of a descending phrase with the words "Duck on my house," repeated at different pitch levels. Pitch is controlled via melodic contours, but there seems to be no overall pitch organization such as we expect in adult songs. The song does not stay in one key, nor does it modulate coherently. Rhythm is relatively well controlled and regular. (It is interesting to compare these songs with the organization of the opening of Beethoven's Fifth—the formal idea is the same in that a contour is repeated over and over, but Beethoven exerts tight control over numerous other parameters.)

When we attempt to teach a child of 2 or 3 a simple song, they usually succeed in reproducing only a phrase or two. Davidson, McKernon, and Gardner (1981) taught a simple song to children spanning the preschool years in age. Children of 2 and 3 were able to reproduce the contour and rhythm of single phrases, but with varying interval size and wandering pitch. As children got older, they were able to combine more phrases into closer and closer approximations of the model. The interval relationships of the adult major scale began to emerge, but only locally within phrases. Four-year-olds could maintain stable pitch and intervals for a phrase or two but then slid to a new key in the next phrase. Their reproductions of the whole song pattern from beginning to end were fairly good.

The progress of the child from 2 to 4 in song learning is strikingly parallel to progress in learning stories. At the earlier age, the child tends to focus on an isolated incident from the story, for example, "Bad wolf chase pig." This is repeated, just as one phrase of a song is repeated. Later, more and more incidents are integrated into a meaningful sequence with the beginnings of a coherent plot. At the later ages, the child integrates several different phrases into a coherent song, both with spontaneous songs (Dowling, 1984) and familiar nursery songs. The child of 3 can sing songs as elaborate as the "Alphabet Song" ("A, B, C, D, . . . " to the tune of "Twinkle, Twinkle") through from top to bottom, getting the correct order of words with the correct melodic contours, but of course
5q with wandering pitch. (Example 5q).

Around the age of 5 or 6, the child acquires the sense of a stable key. Two converging lines of evidence lead us to this conclusion. First, Davidson et al. (1981) found that 5-year-olds could reproduce their little song reasonably well, staying within a single key throughout the whole song or through large sections of it. When the key changed, it was generally a sudden shift to a new key that was then stably maintained, rather than by gradual wandering. Second, Bartlett and Dowling (1980, Experiment 4) found that children of 5 or 6 noticed changes of key when the change was

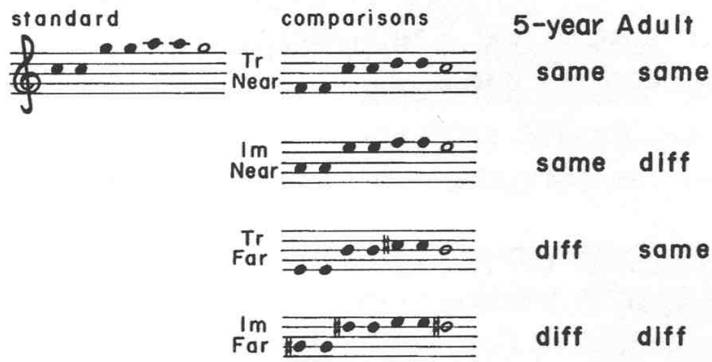


**Figure 5.10** Schematization of results of Bartlett and Dowling (1980, Experiment 4). Tr, Transposition; Im, Imitation.

to a distant key (i.e., one that introduced several changes in pitches of notes in the scale) but not when the change was to a nearly related key (that introduced few such changes). This result, obtained with short-term recognition of familiar melodies like "Twinkle, Twinkle," was an instance of the key-distance effect discussed above. But the 5-year-old has only one component of the adult behavior pattern, namely, the ability to notice a change of key. The child at this age is still unlikely to notice small changes in the pitch intervals of familiar melodies that are obvious to the adult. Figure 5.10 gives a qualitative outline of this result. The 5-year-old responds on the basis of key but not interval size. Bartlett and Dowling found that by 8 years, the child had generally developed the adult pattern, responding to changes of both key and intervals.

The child's development of the ability to notice changes in the key of a melody is closely related to other aspects of melodic information processing. It is around this age that the child becomes able to utilize the fact that a tone sequence is tonal (rather than atonal) in order to detect changes in its intervals. (Here we refer to more noticeable changes than those in Bartlett and Dowling, 1980.) Zenatti (1969) gave children the rather difficult task of saying *which* note of a three-note melody had been altered in pitch. Five-year-olds could not perform this task any better than chance, but by 6 or 7, performance was better than chance, with performance on tonal sequences markedly better than performance on atonal ones. The 6-year-old can use an internalized schema of the tonal scale in performing the task. Another related phenomenon is Imberty's (1969) finding that by 7 years, children could notice sudden changes of key in the midst of a tune. By 8 years, the children in Imberty's study could notice changes of mode from major to minor—a change that produces essentially the same kind of changes in intervals as those in the imitations in Figure 5.10. Here again, the 8-year-olds showed the adult ability to notice changes of both key and interval size. This developmental pattern has been further corroborated by Krumhansl and Keil (1982) with methods akin to those of Krumhansl described in Chapter 4.

## Training

This review of development has so far been generally true of children with little or no specific training in music. Virtually all of the studies with adults we have mentioned found some improvement in task performance with increased musical training. In most cases, both the amount of training and the performance increment were modest. Further, we have known for some time that adults exposed to special training programs in pitch discrimination and melody recognition improve their test performance for those abilities (Wyatt, 1945). The question remains of how young a child can be and still benefit from training. We believe with the Piagetians that a child can only be effectively taught a skill when ready—when the underlying cognitive capacities are sufficiently developed to learn it. However, we also believe that the child is more ready at even the earliest ages to learn musical skills than our culture typically expects. This should be clear from the above review. Who would have expected the success of Kessen et al. (1979) in teaching infants of 6 months to match pitches? We believe that 3- and 4-year-olds can benefit substantially from musical training suited to their abilities but, unfortunately, have little in the way of solid evidence. We do have good evidence that 6-year-olds develop their skills rapidly with training, and to that we now turn.

Dowling and Goedecke (in preparation) studied first and third graders in inner-city public schools who had been enrolled for about 6 months in a training program in either strings or piano. The training programs were based upon the methods of Shinichi Suzuki (1973); that is, they emphasized auditory processing skills and progressive control over the sounds made by the instrument rather than, for example, learning to read music. Children who had had this training during the year were compared with children in a control group drawn from a waiting list of those who desired training but who, for budgetary reasons, could not receive it. Dowling and Goedecke used two short-term recognition memory tasks involving novel five-note tonal melodies similar to those used in Dowling's studies with adults. In the first task (repetition) the child was asked to distinguish between exact repetitions of melodies and lures having different contours and pitches. As could be expected from the performance of Chang and Trehub's (1977a) infants in recognizing contour changes, even untrained first graders did well on this task, achieving about 75% correct (where chance was 50%). In the second task (permutation), listeners had to distinguish between exact repetitions of melodies and lures in which the pitches of the standard were presented in permuted order. That is, lures in the permutation task differed from targets in contour but not in their component pitches. This task was much more difficult than the first, and neither first nor third graders in the control group attained better than 70% correct.

Figure 5.11 summarizes the performance of the various groups of children. It is clear that even first graders benefit from training in auditory information-processing skills. This can be seen in their improvement on the repetition task from 75 to 90% correct. By third grade, performance on this relatively easy task appears to be approaching asymptote for both trained and untrained groups. Training seems to have little impact on performance of the more difficult permutation task for the first graders, but definite gains can be seen in the performance of the trained third graders. Not that this task, unlike the repetition task, is one that benefits little from simple maturation during this period. Untrained third graders do little better than untrained first graders. Nevertheless, third graders acquire the skill effectively with training. What they have learned is to ignore an obvious, salient source of similarity between targets and lures

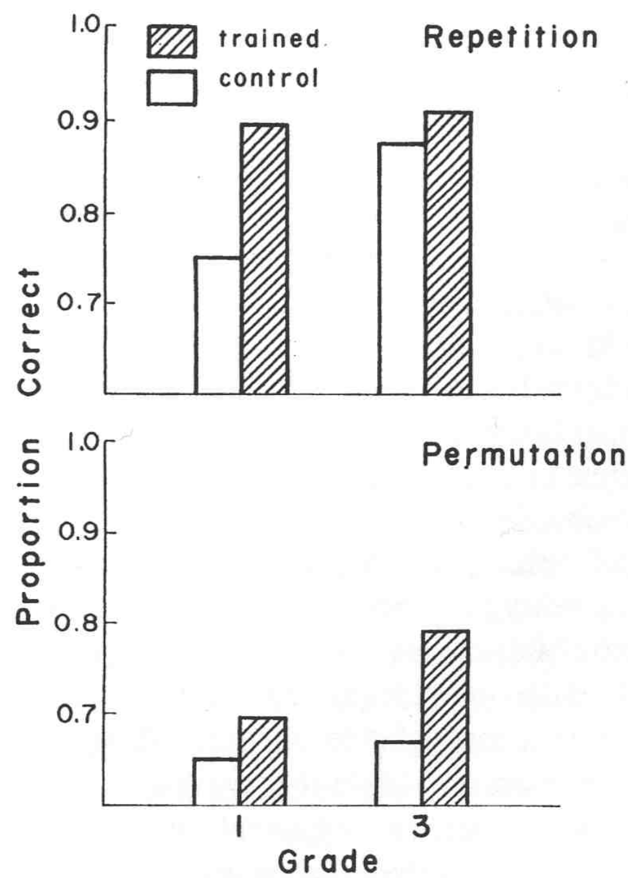


**Figure 5.11** Performance of first and third graders with and without training on the two tasks in Dowling and Goedecke's (in preparation) study.

(identity of pitch content) and focus on the musically meaningful feature of melodic contour. The data of Figure 5.11 show that (1) first graders are cognitively ready to benefit from musical training in terms of improvement in their auditory processing skills, and (2) first and third graders are ready to benefit in different ways, as the review of their typical capabilites suggested.

## SUMMARY

Melodies seem to be listened to and remembered according to a few perceptually salient features. Mental schemas, developed from early childhood in the course of hearing many of the culture's melodies, search for and extract these features from novel melodic information. These schemas include more general, higher-order information than specific pitches at specific temporal intervals. The context in which a pitch is heard affects memory for that pitch, suggesting that more global features are important. Melodic contour, for example, facilitates the immediate recognition of melodies; however, the relative importance of contour varies with tonal scale context. The most important of these contexts is the key-distance effect, which suggests that intervals—patterns of relative pitch chromas—are important in long-term memory for melodies, where a particular melody must be differentiated from a large number of similar alternatives. We argue that task demands affect which melodic features are important in attending to and remembering melodies. Also, differences resulting from prior musical training point to the role of individual differences in a theoretical model of music cognition.

The child's developing schemas for auditory cognition also involve the use of contour, interval sizes and chromas, and tonal scale system as important features. Infants can recognize changes in melodic contour and can produce a match to a given individual pitch. Over the first 2 years, such behaviors become increasingly integrated into coherently organized songs, although still without stable pitch levels or intervals. By age 5 or 6 a sense of stable key emerges, and children can recognize a change in key. We also argue that children are more ready at even the earliest age to learn musical skills than our culture typically expects.

Having established the importance of schemata in the development and use of melodic perception and memory, we turn, in the next chapter, to the larger perspective of melodic organization—to the temporal patterning of melodies, where cognitive schemas become the foundation for memory and comprehension of more complex musical structures.