From Siu-Lan Tan, Peter Pfordresher, and Rom Harré, *Psychology of Music: From Sound to Significance,* New York: Taylor & Francis, 2010.

5 Perception of musical pitch and melody

The following case of a 61-year-old Canadian man was documented by Peretz (1993):

> GL was referred to us in 1989 because of his persistent amusia . . . GL stated that he was totally unable to pick out familiar music and did not enjoy music any more. His musical background is that of a non-musician who was nonetheless an avid listener to popular and classical music, attending concerts and musical recitals very regularly before his illness . . . Contrasting with his relatively good linguistic abilities, recognition of tunes (without lyrics) was totally abolished. Out of 140 musical excerpts (including his national anthem) that are very familiar to everybody in Quebec, he could not identify a single one. (p. 27)

About 10 years earlier, the man had suffered an aneurysm on the right side of the brain followed a year later by a mirror aneurysm on the left side, after which he was diagnosed with severe *Wernicke's aphasia* and *amusia* (broadly defined as loss of language comprehension and music ability, respectively). He recovered most of his speech abilities but the amusia persisted. When tested on his ability to discriminate isolated pitches, GL's performance was comparable to matched controls. However, when pitches were presented in a melodic context, clear deficits were seen. For instance, GL had difficulty accurately identifying whether two melodies that were altered by one tone were the 'same' or 'different.' He was also unable to correctly identify whether a melody sounded 'complete' (ending on the tonic tone) or 'incomplete' (ending on a nontonic tone). As described above, he could not identify well-known melodies such as *O Canada*, his own national anthem! Peretz concluded that while processing of isolated pitches was intact, GL had 'lost access to tonal knowledge' (1993, p. 51),[1] which – as we shall see in this chapter – is absolutely fundamental to 'making sense' of melody.

The case of GL presents us with some intriguing questions about the perception of melody, a basic musical ability that most of us take for granted. How do we learn and retain the thousands of tunes we 'know'? Melodies are simply different arrangements of discrete pitches with distinctive rhythmic patterns. How are we able to remember each particular arrangement? If melodies are made up of separate pitches, what binds a melody together so that it is heard as an organized and coherent tune? These questions serve as the starting point for our discussion on melody and pitch. We will review the classic work on the perception of Gestalt wholes, explore some findings on memory for pitch and melody, and conclude with related studies in the neuroscientific research.

The ingredients of melody

In keeping with standard definitions of melody (e.g., Radocy & Boyle, 2003), we consider the term 'melody' to characterize one aspect of musical experience: the experience of a sequence of pitches as belonging together. This definition captures the idea that melodies are perceived, not in terms of their separate constituent tones, but as coherent units. While each tone of a melody reaches the listeners' ears as if it were a single bead, listeners 'thread' the beads together into continuous strands. We will turn to the topic of the perceived unity and coherence of a melody line in our discussion on Gestalt principles of perception. But first we will explore some essential elements of melody: pitch, interval, contour, harmony, and key.

Pitch

We already considered the concept of pitch at a psychoacoustic level in chapters 2 and 3. However, as we have seen in the case of GL, when pitches are perceived within the context of a musical melody our perception of pitch becomes more complex. It has been pointed out that musical pitch is a *multidimensional* percept; that is, there are multiple continua along which pitches may be distinguished (Shepard, 1982). We focus on two here. One dimension, true of pitch either in or out of a musical context, is *pitch height*, related to the frequency of vibration. However, a more musically important dimension of pitch is its so-called *chroma*. Chroma refers to the category (or 'class') represented by a certain pitch. The names we give notes in Western tonal music (e.g., C, D, E) refer to pitch chromas. Chromas are identical when separated by an octave (a 2:1 ratio of a tone's fundamental frequency) and thus constitute a distinct dimension from pitch height. The fact that tones separated by an octave inhabit the same chroma category is called *octave equivalence*. Within an octave, however, changes in chroma are also changes in pitch height; the dimensions are therefore distinct but not independent. Evidence that the auditory cortex maintains distinct representations for pitch chroma and pitch height has been reported in an fMRI study (Warren, Uppenkamp, Patterson, & Griffiths, 2003).

So far we have established that one way to characterize melodies is as a sequence of tones, with each separate tone containing particular information about pitch height and pitch chroma. However, this is probably not the way most people conceptualize melodies. Consider the fact that most individuals, even from an early age (i.e., infancy, as discussed in chapter 8), perceive a melody to be the same when it is played in a different key (a 'transposition'). If 'Over the Rainbow' is transposed from the key of C to D major, the opening phrase changes from C-C′-B-G-A-B-C′ to D-D′-C♯-A-B-C♯-D′, which changes every pitch in the sequence. Yet most people have no trouble recognizing these melodies as the 'same.' In transposed melodies, even though specific pitches may be changed, the *relationships (i.e., intervals)* between the pitches are retained, and it is these relationships that people primarily use when recognizing a melody. We now turn to ways in which the relationships between pitches may be characterized.

Interval

Intervals denote the transition from one pitch to the next. In musical practice, one more commonly hears of scale step intervals (e.g., major third, perfect fifth), which assume a tonal structure. The interval level of description is an abstraction because it loses information about specific pitches. Its significance to the cognitive representation of music is that people can typically recognize and reproduce a melody in many different keys. This accomplishment relies on the fact that people represent melodic structure in a way that is independent of pitch class. This is what is meant by the term *relative pitch* perception, the ability to recognize and remember pitches by virtue of their relationships to one another (i.e., intervals between pitches) as opposed to the specific identity of individual pitches. Individuals who possess *absolute pitch* (or AP, commonly referred to as *'perfect pitch'* among musicians) can identify the pitch chroma of isolated tones (passive AP) or produce a chroma (active AP), without being given a reference pitch. Thus AP possessors do not rely on intervallic relationships to recognize pitches. We shall return to the fascinating topic of absolute pitch later in this chapter.

Melodic contour

Melodic contour refers to the shape of a melody line, depending on whether successive pitches are rising, falling, or unchanging in pitch. Like the distinctive skyline of a city, it is the shape or outline that gives a melody its distinguishing character. Researchers are interested in melodic contour because changes in direction constitute salient points in a melody (e.g., Jones, 1987). Moreover, a performance error that results in a change to the melody's contour is far more noticeable (and far less common) than an error that does not cause such a change. GL was unusual as he had difficulty discriminating between short melodies even if the alteration modified the direction of the pitch, changing the overall contour (Peretz, 1993). However, as we will see in chapter 8, even infants can discriminate between two melodies that differ only in one pitch if it alters the melodic contour (e.g., Trehub, Thorpe, & Morrongiello, 1985).

Harmony

Harmony and melody are often referred to as if they were independent, so it is perhaps odd to see harmony listed as an ingredient of melody. However, the harmonic scheme can influence the perception of melody in many ways. For example, the harmonic structure can imply segmentation of the melody into sections by virtue of harmonic changes. When a phrase ends with an authentic cadence (harmonic progression from dominant to tonic, or chord V to I), the end of a section is strongly implied, even if the music continues after that cadence. These boundaries form a higher level of structure than do changes in interval and contour. Melody can also imply harmony even in the absence of any accompaniment. A melodic sequence of notes such as [G A G E C] gives the strong implication that if one were to play a chord along with these notes that the chord ought to be a C major tonic triad. Indeed, our ability to recognize melodies depends in part on our sense of implied harmony (Cuddy, Cohen, & Mewhort, 1981), especially for musically trained listeners (Schubert & Stevens, 2006). We have assumed diatonic scales (e.g., major, minor) in our discussion, but there are many other scale and tonal systems around the world (see chapter 15).

Key

Key refers to the kind of scale structure that a melody implies. In the diatonic scale system it can be *tonal* or *atonal*. Tonal melodies are those for which the constituent pitches imply a specific key, whereas for atonal melodies there is no such implication. If tonal, the key can be major or minor, or another mode. Key may constitute the most abstract way to describe a melody, in that a melody can often be referred to with respect to just a single key (although key modulations are common, which relegates key to a more local level). As a description of melodic structure, however, key is highly limited. There is no way to recover the structure of a melody if all one stores in memory is the key. Key places important limitations on notes that sound as if they 'belong' to a melody. As with contour, errors that violate key are highly salient. Similarly, out-of-key notes that are consciously added to a composition can create moments of high tension. GL was a rare exception; when asked to select a tone to complete a five-tone sequence, he preferred tones that were not part of the scale to diatonic tones.

Not all pitches are 'equal' within a scale. For instance, the *tonic* and *dominant* (first and fifth degrees of any scale) are structurally important, and usually give a sense of a stable center and closure (tonic) or repose (dominant). Indeed, music theorists view every tone within the diatonic scale as having a function in relationship to the other tones (as indicated by their technical names such as 'tonic,' 'dominant,' 'leading tone,' or 'leading note'). Harmonically, the *tonic*, *dominant*, and *subdominant chords* (chords I, V, and IV) in any key are structurally most important. Many simple songs can be harmonized simply by using just these three chords. Simply through exposure to music, even listeners with absolutely no musical training possess some general knowledge structures of tonal music that are activated when listening to music, enabling them to form basic expectations about where the next note or phrase might go, or if a song has come to a 'good finish.' GL represents a very rare illustration of how, in the absence of tonal knowledge providing the context, melodies lack meaning. While explanations for the structural importance of some tones of the scale over others are provided by music theorists and physicists, we will focus on a study exploring the perceptual bases in our discussion of 'tonal schemata' and Krumhansl's (1990) research later in this chapter.

The perceptual organization of melodies

Earlier, we compared the separate tones of music to loose beads, which listeners thread together into strands. When listening to more complex music beyond the single melody line, a more complicated network of these strands of beads emerges. In many cases, there may be many different ways that the beads may be organized, as a complex work is open to many interpretations. When listening to a pop song, jazz standard, or symphonic orchestral work, there is a dynamic interplay of melody and harmony lines, and relationships between voices and instrumental parts. In part, it is this ability to organize incoming musical input into coherent units that makes music listening such a rich and enjoyable experience. In this section, we take a closer look at how it is that listeners organize discrete musical sounds into coherent wholes as a composition unfolds, focusing on the perception of melody.

Gestalt principles of perception

The word *Gestalt* is German for 'whole form.' The study of the conditions under which 'gestalts' are perceived in a sequence of elementary parts is the main aspect of Gestalt psychology that is of interest to us. In other words, gestalts are the organized structures that emerge from the physical stimuli in our environment.

This development was initiated by Max Wertheimer in Berlin before the First World War and extended by others, notably Kurt Koffka (1896–1941) and Wölfgang Köhler (1887–1967). The Gestalt pioneers began systematic studies of how the elements in an aggregate must be arranged in order for it to display an emergent property, to be seen or heard as a whole. According to the Gestaltists, the perception of the emergent or whole property of an aggregate and of the elements contributing to that whole influence each other. We see patterns in groups of elements, and the perceived groups in turn influence our perception of the elements (Wertheimer, 1938).

The following are examples of basic Gestalt principles describing how elements group to form an aggregate. These principles, originally designed to be applied to vision, follow from a single superordinate principle called *Prägnanz*, which dictates that perceptual organization conforms to a form that is simple and symmetric (Koffka, 1935, p. 110):

> *Proximity*: Other things being equal, the elements that are near to one another tend to be seen as a group.
>
> *Similarity*: When more than one kind of element is present those which are similar tend to be seen as a group.
>
> *Closure*: When a pattern is incomplete, there is a tendency to perceive it as whole and complete by 'closing' the gap.
>
> *Good continuation*: Smooth continuity is preferred over abrupt changes of direction.

Although first applied to visual phenomena (as illustrated in Figure 5.1), these principles have a more or less direct application to the perception of musical wholes, the components of which are tones.

Gestalt principles and music

In its application to music, the *Principle of Proximity* may explain how a series of tones is perceived as a melodic line, as opposed to a series of unrelated and disconnected tones. In music, 'nearness' may be applied in several forms. Tones that are close together in *pitch*, *time*, or *space* tend to be perceived as a group. For instance, tones separated by large leaps in interval, separated by rests or pauses, or originating from different spatial locations (e.g., music coming from left versus right loudspeakers) tend to serve as boundaries between groups. Of these, proximity in pitch appears to be one of the most important factors in governing grouping of auditory streams. Melodies in both Western and non-Western music tend to be comprised of small intervals and relatively narrow pitch range (e.g., see Dowling, 1968).

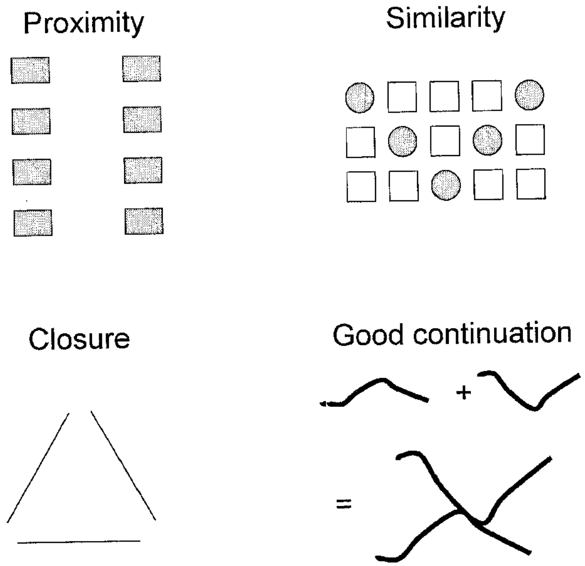

Figure 5.1 Examples of Gestalt principles, with visual examples. Copyright © Peter Pfordresher.

Indeed, Diana Deutsch has demonstrated how 'octave-scrambled' presentations of familiar melodies are difficult to recognize. Although most effectively demonstrated in Deutsch's CD *Musical Illusions and Paradoxes* (Deutsch, 1995, Track 19–23 'mysterious melody'), the effect may be easily demonstrated on a piano keyboard by playing a well-known melody on a piano in an 'octave-scrambled' fashion. Simply preserve the original notes and rhythms but play each note in a different octave (most effective if using full range of the keyboard, and selecting octaves without any pattern). Not only will most listeners be unable to identify the 'octave-scrambled' melody, but most will also not experience the sequence as a coherent 'melody.' Due to the gross violation of the Principle of Proximity, the tones do not cohere into a melody but seem to 'fly apart'!

The same principle is at play in a classic study by W. Jay Dowling (1973). Dowling investigated the conditions under which listeners are able to separate two 'interleaved' (or intertwined) melodies. Interleaving was accomplished by playing alternate tones between two familiar melodies (i.e., first tone of melody A, first tone of melody B, second tone of melody A, second tone of melody B, and so on). This musical device is often used by composers and orchestrators. Here is a visual analogy of interleaving, employing interwoven sequences of letters.

INTERLEAVED

TWOMELODIES

When interleaved, the result would be:

ITNWTOEMRELLEOADVIEEDS

Dowling showed that when two familiar melodies are interleaved, they are difficult to distinguish when the melodies presented are overlapping in pitch ranges. When played in different pitch ranges, however, the two groups of tones are heard as two distinct melody lines. Many examples of this effect can be found in musical works, in which alternating between high and low registers may create the impression of two melodic lines played by a single instrument (*virtual polyphony*) as exemplified by Bach's solo violin works (e.g., Davis, 2006), or in single voice as in yodeling. Due to our tendency to group pitches by proximity, two melodic lines are heard, as opposed to a single jagged melody alternating between high and low tones.

The *Principle of Similarity* also plays a significant role in the perception of music. There are times when we attend to a single voice or instrument. But for the most part, whether listening to a local garage band or a symphonic orchestra, we organize the input of individual sounds into organized groups such as 'melody' and 'accompaniment,' 'strings,' 'winds,' 'brass,' and so on. In such circumstances, it is often timbre that serves as the point of similarity that allows the listener to segment sounds into different instrument groups.

Another application of the Principle of Similarity in music is sequential similarity, by which repetition and variation of themes, motifs, and relationships among musical ideas facilitate our perception of a musical work as a coherent whole. The elements in a visual array are often simultaneously present for inspection of repetition or patterns, but many forms of musical similarity unfold sequentially, in time. Ravel's *Bolero* is a marvelous study in repetition, with its insistent theme repeated throughout the piece over an ostinato rhythm. Sometimes, as in the child's round 'Row Row Row Your Boat' or Eric Clapton's 'Layla' (which opens with six exact repetitions of a riff) the repetition is quite apparent to a perceptive listener. In other cases, such as the doubling of the theme in other keys as the orchestra augments in *Bolero*, or appearances of the subject in a complex three- or four-part Bach fugue, the relationships may be more difficult for the listener to perceive. It is our ability to hear repetition, patterns, and other musical relationships that lends a sense of unity and order to a work of music as a whole.

It should be noted, however, that these principles of perceptual coherence may hold for short simple melodies but perhaps not for more extended works (see Levinson, 1997). For instance, when Tan and Spackman (2005) created 'patchwork' compositions by linking extracts of music by different composers (e.g., Schumann–Liszt–Chopin) together with abrupt changes in pitch, key, harmony, tempo, style and little repetition of melodic ideas, few (musically trained or untrained) listeners noticed that the music had been structurally altered. In another study, Tan and her colleagues found that after four repeated hearings, listeners even came to prefer the 'patchwork' compositions to intact compositions, and rated them higher in musical unity (Tan, Spackman, & Peaslee, 2007). In a study in which 'hybrid' compositions were created from combining pieces of different Mozart piano sonatas, Eitan and Granot (2008) also found that with repeated exposure, listeners showed a preference for the 'hybrids' and, surprisingly, musically trained participants preferred the hybrids more than untrained participants. It is important to keep in mind that Gestalt principles that may easily be demonstrated with short melodic sequences may not always apply to more complex and extensive musical works.

The *Principle of Closure* is exemplified in the way that finality as completion of a melody is expressed by resolution to the tonic of the scale, either in the melodic line or in the actual or implied harmonic progression. We are reminded of the almost surely apocryphal story of how a visitor got Mendelssohn out of the bath. As the story goes, a visitor was informed that Herr Mendelssohn was in the bath and could not see him. The guest went to the piano and repeatedly played the first seven notes of the scale of C major. After several repetitions, Mendelssohn suddenly appeared in his bath robe – and completed the last note of the scale![2] The leading tone (or seventh note of the diatonic scale) sets up such a strong expectation to resolve to the tonic that most listeners imagine the last note, 'closing' the scale with a return to the tonic – much like one mentally fills the gap to close an incomplete circle. In Mendelssohn's case, however, it was not enough to mentally complete the scale. According to the story, he was absolutely compelled to *play* that tonic note! Such strong expectations based on the functions of the tones in a scale were mentioned earlier; pitches in a melody are associated not only with a current value but with a kind of musical velocity. In the case of GL, who did not seem to hear the pitches in a tonal context, the principle of closure was not at play; he did not seem to feel the 'pull' to the tonic.

Experimental evidence for closure in music was reported by DeWitt and Samuel (1990). The authors tested whether an effect similar to phonemic restoration in language occurs in music. *Phonemic restoration* (Warren, 1970) is a powerful effect in which listeners continue to hear intact speech even when a phoneme is replaced by noise. Rather than hearing the noise where it occurs, the listener experiences the speech sound as if continuing *through* the noise. This effect is analogous to 'filling in a gap' visually, and tends to be stronger at the end of sentences than at the beginning (suggesting that the process of filling in requires knowledge of the context, which is referred to as 'top-down processing'). In their investigation of *perceptual restoration* in music, DeWitt and Samuel (1990,[3] expt. 4) created 10-tone scales (octave scale plus next two tones) with one tone either present but accompanied by noise or replaced by noise. Listeners had greater difficulty discriminating whether the tone was present or absent with the noise if the distorted tone was the expected pitch that would complete the scale (as opposed to a random pitch). They were also more likely to 'restore' a missing pitch as the number of tones of the scale before the distorted tone increased. Thus the findings showed 'an increase in restoration associated with increased musical expectations' (p. 141).

While Gestalt principles account for our natural tendency to group musical sounds in organized ways, they also explain the principles by which particular auditory streams of musical sounds are heard as distinct from the rest. In particular, this pertains to the question of how a melody line (the 'figure') can be heard against sometimes very complex and busy accompaniments ('ground'). For a thorough discussion of auditory stream segregation in instrumental and vocal ensembles, the reader is referred to Albert Bregman's *Auditory Scene Analysis* (1990), especially chapter 5 and pages 490–502.

In the first chapter to this important book, Bregman states that 'the Gestalt principles of grouping can be interpreted as rules for [auditory] scene analysis' (p. 24).

The scale illusion

In most cases, our natural tendency to sort incoming sounds in organized ways helps us to make sense of complex auditory stimuli. Under certain conditions, however, the Gestalt principles may be in play where in fact they should not apply and this can lead to false impressions about what we are actually hearing. One of several interesting auditory illusions described by Diana Deutsch (1975), and used here as an illustration, is the *scale illusion.* In Deutsch's study, participants listened to a musical sequence through stereophonic headphones. The sequence was based on the C major scale, with the notes of the ascending and descending scale alternated simultaneously between the left and right ears, producing two rather jagged melodic lines played to each ear (as illustrated schematically in the top part of Figure 5.2).

Participants were asked to give a verbal report of what they heard through the headphones, and to sing what they heard in each ear. Not one of the 70 participants reported hearing the musical pattern that was actually played! As illustrated in Figure 5.2, most (right-handed) participants 'heard' smooth ascending and descending contours of the scale, with the higher tones in the right ear. The most common percept was half a descending C major scale which then ascended back to the C in the right ear, and half an ascending C major scale in the left ear which then descended back to the C. Thus, the tones not only 'reorganized' into smooth melodic contours, but seemed to 'migrate' from one ear to another to do so. When (right-handed) participants were asked to reverse their stereophonic headphones so that the input formerly presented to the left ear was presented to the right ear and vice versa, most participants' reports did not change. Again, they heard a smooth descending-ascending scale pattern in the right ear, and the mirrored contour (all the lower tones) in the left ear. The responses of left-handed participants showed more variation, and no clear pattern was reported for them.

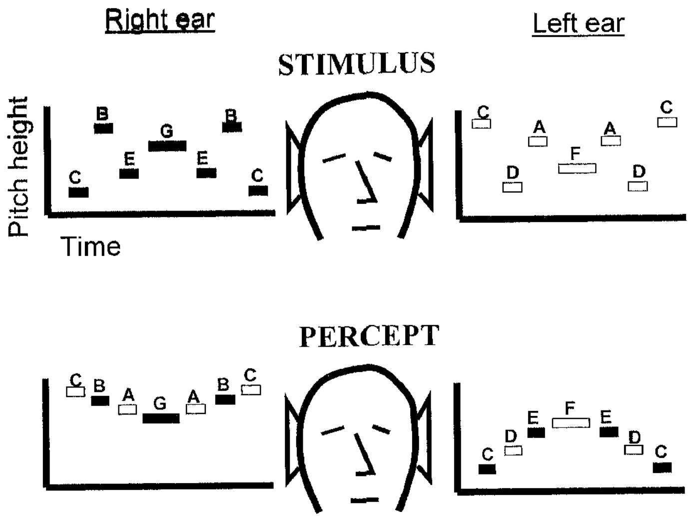

Figure 5.2 An illustration of the scale illusion. Rectangles denote tones varying in pitch and timing. Filled rectangles denote tones presented to the right ear (through headphones) and open rectangles denote tones presented to the left ear. Adapted from Deutsch (1999).

The 'scale illusion' illustrates several Gestalt principles of grouping, especially the *Principles of Proximity and Similarity*. If the tones are played in sufficiently different pitch ranges, sufficiently different timbres, or if the tone sequence does not imply an orderly scale, the effect is likely to be weakened or absent, as the perceptual principles do not act singly but often act 'in concert' or in competition with one another. For instance, it is interesting that most listeners do not hear a full descending scale in one ear and a full ascending scale in the other, analogous to our visual illustration of the *Principle of Good Continuation* in Figure 5.1. Other predispositions, such as the strength of the proximity principle and our tendency to more accurately localize high sounds in the right ear and low sounds in the left ear, may override the tendency to sort the tones into continuous lines that cross over each other.[4]

Memory for pitch and melody

Having explored the immediate perception of melody, we now turn our attention to memory for melodic information. In this section we discuss memory for pitch, contour, tonality, and an example of implicit memory for music.

Memory for pitch

It should be noted that the aforementioned studies assume that melodies are remembered without reference to specific pitches, as described earlier. This limitation makes sense given the common assumption that few people encode melodies using absolute pitch. The ability to label individual pitches that one hears without being provided with a reference pitch may be fairly rare, possibly occurring in only 1 of 10,000 persons (Takeuchi & Hulse, 1993). Perhaps in part because of its rarity, *absolute pitch* (or *perfect pitch*) is a topic of widespread fascination, in spite of the fact that relative pitch may be a more useful ability for music perception, due to the relative nature of melody. For instance, musicians with relative pitch outperformed those with absolute pitch when asked to judge whether two transposed melodies preserved intervallic relationships or not (Miyazaki, 2004). In comparison to absolute pitch, the ability to recognize pitch intervals is more common and it is widely assumed that such a *relative pitch* representation dominates the processing of musical pitch for most persons.

Recent evidence, however, suggests that most people in the general population may store some melodies, not just as abstract relational representations, but as absolute pitch representations in long-term memory. An early study demonstrating this was run by Levitin (1994) who simply asked participants to sing their favorite tune from memory, beginning wherever they wish. Surprisingly, these participants (who were not trained singers and did not possess absolute pitch) produced melodies within 1 or 2 semitones of the original key of the melody most of the time. Importantly, the songs used in Levitin's study were recorded pop tunes that listeners had always heard in the same key. This study showed that although correct *pitch-labeling* (C, D, E♭, etc.) without a reference pitch may be rare, good *pitch memory* seems to be widespread. Similarly, Bergeson and Trehub (2002) found that mothers singing songs to infants were very consistent in pitch and tempo when recordings of singing (separated by intervals of one week or more) were compared.

A plausible explanation – according to one view – would be that participants transpose melodies into the key most comfortable to their singing range, resulting in a fairly uniform distribution of produced keys. However, other studies showed that evidence for accurate pitch memory is not limited to production. For instance, Schellenberg and Trehub (2003) played orchestral theme songs from popular television programs (including *Friends*, *E.R.*, and *The Simpsons*) either in their original pitches, or shifted up or down by only one or two semitones. Most people were able to identify the original key of familiar television theme songs despite the variations being very close in pitch, leading the authors to conclude that 'good pitch memory is widespread' (p. 262). Other studies have shown that people can identify tunes from MIDI (Musical Instrument Digital Interface) recordings within one phrase (Dalla Bella, Peretz, & Aronoff, 2003), and can identify audio recordings within 200 milliseconds (Schellenberg, Iverson, & McKinnon, 1999).

Does pitch perception change with age? One of your authors (ST) was identified as having AP at around age 7, while being tested by an examiner during an aural exam for the Associated Board of the Royal Schools of Music. She was quick at identifying single pitches and could even name pitches comprising three- or four-note chords with high accuracy. During her mid-thirties, she began to make more errors when identifying pitches (especially in the extreme high and low ranges) and tended to be a semitone sharp when incorrect (e.g., hearing the lowest E♭ on a grand piano as an E). She has since discovered that this 'shift' is common. Athos and colleagues (2007) tested almost 1000 AP possessors between the ages of 8 and 70, and found that 'pitch errors [among individuals with AP] increase with age, and they tend to be sharp' (p. 14796). In fact, one 44-year-old participant did not identify a single pitch correctly during the test but was consistently one semitone sharp, a tendency he had observed since age 22. The reason for this shift is not yet known but Athos and colleagues (p. 14797) surmise that it may correspond with some 'age-dependent physiological change that alters the mechanical properties of the cochlea' (whose critical role in hearing was discussed in chapter 3). Indeed, the authors hypothesize that 'such a gradual perceptual shift is common to most people as they age, yet they are unaware of it unless they have AP' (p. 14797).

Memory for melodic contours and intervals

We mentioned earlier that melodic contour constitutes the most basic and possibly the most salient attribute of melody. As it turns out, the influence of melodic contour on memory is strong but it is also qualified. It appears that the only kind of melody for which melodic contour entirely determines memory confusions may be an *atonal* melody. When melodies are *tonal*, key also contributes to melody recognition. Dowling (1978) reported a study in which people heard pairs of melodies and had to determine whether the second melody constituted a transposed version of the first melody or was a different melody. Whereas musically trained listeners were good at ruling out an atonal melody that preserved the original melody's contour, they were not as good at ruling out 'tonal answers,' which preserved contour and key but differed with respect to constituent intervals. Tonal answers are akin to repeating a melodic theme in different registers of a common key. For instance, if one plays [C E D F G] followed by [F A G B C] the two excerpts sound quite similar by virtue of their common contour and their similar reference to the key of C, even though the last two intervals differ. Unskilled listeners are more likely to be fooled by atonal melodies that match in contour, though they too are better at ruling these melodies out than tonal answers.

The results summarized in the previous paragraph apply to melodies held in memory over the short term. As we become increasingly familiar with melodies, our memory starts to store interval information as well, and as a result people get better at ruling out different kinds of alterations. It is highly unlikely, for instance, that one would be fooled by a 'tonal answer' to 'Happy Birthday.' This finding was originally reported by Dowling and Bartlett (1981; see also Dowling, Kwak & Andrews, 1995).

Tonal schemata

Musical memory need not be specific to a particular melody – it can work in more general ways. One uncanny aspect of music perception that we often take for granted is the fact that a 'wrong note' can really stand out perceptually, even when the wrong note is very close, physically speaking (for example, in spatial location on a keyboard), to the 'correct' note. And 'nearby' wrong notes (C♯ instead of C when the key is C major) can be more noticeable than some 'distant' wrong notes (such as playing a G instead of a C when the key is C major). A broader implication of this fact is that tonal musical contexts cause listeners to categorize tones in ways that often contradict basic psychoacoustic aspects of pitch – perceptual and physical 'closeness' are not necessarily the same thing.

The implication of this research is that listeners maintain internalized rule structures (commonly referred to as *schemata*) in long-term memory and use these rule structures to interpret incoming sounds such as melodic sequences. Tonal schemata differ from those discussed earlier for melody in that tonal schemata are typically not thought to rely on or to use sequential information. Rather, tonal schemata are used to interpret the role that constituent pitches play within a melody, with respect to the key that the melody evokes. (Examples were given in our discussion of key, earlier in this chapter.) Though individual differences have been documented with respect to the way in which listeners acquire and use tonal schemata (see Smith, 1997 for a review) recent research suggests that all listeners may gain these kinds of rule systems through exposure, even if they are not conscious of doing so (Tillmann, Bharucha, & Bigand, 2000).

The leading figure in this area of research is Carol Krumhansl, whose work is extensively discussed in the book *Cognitive Foundations of Musical Pitch* (Krumhansl, 1990). Krumhansl introduced the *probe tone technique*. She presented listeners with a short tonal context, which could be a scale (major or minor), a chord, a series of chords, etc. Following this context, listeners would form some kind of categorical judgment on a tone or pair of tones that followed the context (such as rating the similarity between two pitches). Generally speaking, listeners' categorical judgments were most strongly influenced by the status of tones within the *pitch hierarchy* that was established by the context. In a tonal context, the *tonic* (this is the pitch that gives the key its name, e.g. C in the key of C major) is considered to be the 'most stable' tone, followed by the *dominant* (which is G within the key of C major) and the *mediant* (E in the key of C major), other within-key notes (D, F, A, B), and finally notes that are out of the key.

A central finding was one in which listeners categorized the similarity of two successive tones that followed the context. The rating results were displayed using a technique known as *multidimensional scaling*. The idea behind multidimensional scaling is essentially similar to that of a scatterplot. A scatterplot shows the relationship between two variables, X and Y, in Cartesian coordinates (a rectangular grid). The coordinate system is assumed and, within it, a relationship may look 'good' (e.g., suggesting a clearly linear relationship) or 'messy.' By contrast, the coordinate system in multidimensional scaling is flexible and can be altered to characterize the relationship between (or among) variables in an ideal way. Moreover, when analyzing similarity ratings every single pair constitutes a possible dimension (or variable), and thus multidimensional scaling offers a way to simplify the display of results. Thus, similarity ratings may be plotted in two or three dimensions, on a curved or flat surface, and so on. Ultimately the data are displayed in a way that forms some kind of geometric shape, suggesting that the cognitive representation of the relations between pitches are best captured by this shape.

The similarity ratings that Krumhansl collected best conformed to a conical shape, shown in Figure 5.3 (described by the author as a 'slightly idealized' scaling solution). Though this kind of representation may look unfamiliar, what it does is fairly simple: similarities are shown as geometric distances. Note that different points are associated with pitch classes within the C-major scale (though in principle this shape could apply to any major key). Closer points are heard as more similar to each other. Thus, when the tonal context is followed by the pitch pair [C G], a participant would be likely to rate the pair as highly similar, whereas the pair [C D♯] would be rated as highly dissimilar. Note that Krumhansl's results differ from basic psychoacoustics, in that the fundamental frequencies of C and D♯ are closer (when presented in the same octave, as Krumhansl did) to each other than are the frequencies of C and G. Some have suggested that tonal structures like these do follow from a different principle from psychoacoustics, that is the avoidance of dissonances (e.g., Parncutt, 1989). For instance, minor seconds when sounded together lead to high dissonance due to partials that fall within the same critical bandwidth (see chapter 3). On the other hand, a perfect fifth (closer in the geometry of Figure 5.3) is highly consonant.

Two aspects of the geometry are of central importance. First, note that different levels of the pitch hierarchy line up at different vertical levels. The bottom level includes pitches that make up a C-major triad, which is the 'root chord' for C major. The next level higher includes pitches that are in the key of C major but are not part of the triad. These two levels are closer to each other than the second level is to the third level, which includes pitches that do not fit within C major. A second aspect of the geometry concerns the diameter of the semi-circle formed by pitch relationships at each vertical level. The bottom level (the C-major triad) forms a semicircle with a small diameter. This means that all the pitches at this level are heard as fairly

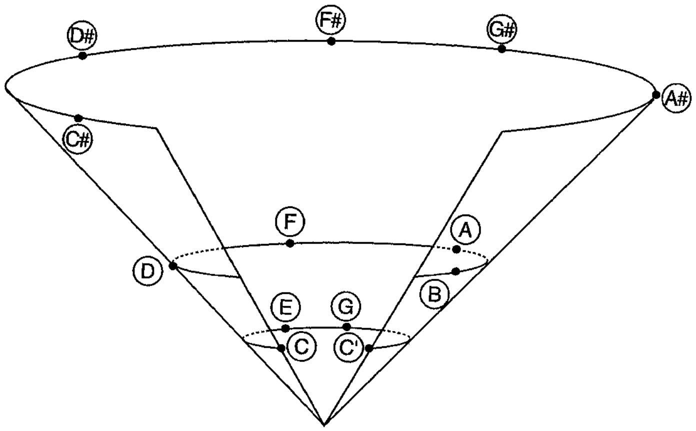

Figure 5.3 Conical representation of similarity ratings for pitches within a C-major context (Krumhansl, 1979, Figure 3).

Source: Reprinted with permission from Elsevier.

similar to each other. In general, you can see that all the within-key pitches were heard as being fairly similar. By contrast, note that the cone widens noticeably at the top level. Pitches outside the tonal context were thus heard as being dissimilar not only to in-key pitches but also dissimilar to each other. The downward orientation of the cone in addition suggests that there is a kind of 'gravitational pull' towards the tonic. With respect to the way we experience music, this model has implications for notes that cause a listener to feel 'tension' or 'relaxation.' Notes near the bottom of the cone sound like endings and thus lead to relaxation, whereas notes near the top suggest that all is not finished, leading the listener to feel a tense kind of suspense.

Note that the schema laid out in Figure 5.3 is true for pitches within a single key. What about relationships among keys? A common music theory perspective is that all the keys within a single mode (major or minor) are related to each other via the *circle of fifths*, which is shown in Figure 5.4. Though the figure shows the circle for major keys, the same layout would hold for minor keys. As with the Krumhansl cone, keys that are closer together around the circle are thought to be more similar. Similarity here, in contrast to pitch, is determined by the set of pitches that make up a key. Keys that are adjacent around the circle have all pitches except one in common; for instance C and G major differ by only one pitch (F versus F♯). By contrast, C and F♯ major keys only have one pitch in common, the rest differ (only B is held in common). Thus, although the tonic pitches for C and D♯ are quite close, the keys associated with these tonics are quite distinct.[5]

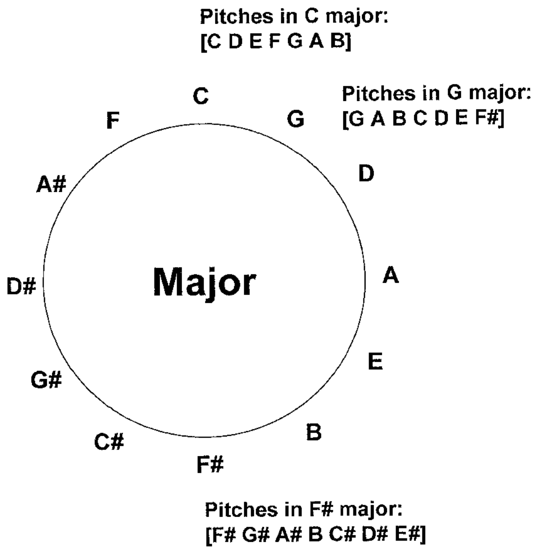

Figure 5.4 The circle of fifths for major scales. Constituent pitches for three selected keys (C, G, and F♯) are given in brackets.

Implicit memory and priming

Memory need not always involve a conscious recollection of past events (explicit memory). This can be illustrated in the role of *priming* in music cognition. The idea is that the perception of a stimulus leads you to access related items from memory. As a result, the perceptual system can more easily encode these related items for a brief period of time. Priming depends on a process psychologists call *implicit memory*. Implicit memories influence how we perform tasks but are not typically accessible to consciousness. In fact it is often difficult to verbalize an implicit memory. For instance we know how to tie our shoes based on memory, but most people find it very difficult to describe how this task is performed!

Many parallels have been drawn between priming effects in language and in music (e.g., Tillmann & Bigand, 2002, for a review). In *semantic priming* in language, the recognition of a word can be influenced by relevant material presented earlier. For example if someone is shown the word 'donkey' the speed with which they can distinguish 'horse' from 'hoarse' is increased (McNamara, 2005). Similarly, in *harmonic priming* in music, processing of a chord tends to be faster and more accurate when preceded by a chord that is harmonically related or 'schematically probable,' compared to one that is not (Justus & Bharucha, 2001; Tillmann & Bigand, 2001). This effect has been demonstrated in studies measuring behavioral responses as well as experiments employing scanning with functional magnetic resonance imaging (fMRI) (Tillmann, Janata, & Bharucha, 2003).

Repetition priming, another kind of priming simply defined as 'a processing benefit for previously encountered stimuli' (p. 693), has been demonstrated with melody (Hutchins & Palmer, 2008). When participants were asked to sing the last pitch of a five-tone melody, they responded faster if the pitch was included as one of the first four tones than if it was a new pitch. Response time was faster when the prime and target were closer together in time, and faster for tonic endings than nontonic endings.

Neural bases of pitch and melody perception

In this final section, we discuss some examples of neuroscientific studies on the perception and memory of musical pitch, key, and melodic contour. This approach is described in chapter 4, and Figure 4.2 may be useful in locating brain regions referred to in the following discussion.

Chroma

One issue that has intrigued researchers in music cognition is the degree to which the brain possesses fixed chroma-specific pitch categories; that is, *absolute pitch*. The dominant view, as discussed earlier, is that most people do not remember 'absolute' pitch, that is, they do not remember specific chromas. From a neurophysiological perspective, the rarity of absolute pitch presents a puzzle. After all, the cochlea of the ear and the auditory cortex share what we call a 'tonotopic' representation of pitch (see chapter 3), which is chroma-specific, similar to a piano. By contrast, we do not have evidence for brain regions that are devoted to specific musical intervals (relative pitch). Why, then, is absolute pitch so rare?

One view, which has received increasing support, is that through the course of development, relative pitch increasingly 'takes over' in humans. Consider the typical musical environment for a human. We usually hear common songs repeatedly such as 'Happy Birthday' sung by different people and in different keys. In order to learn such culturally significant melodies, the listener must recognize similarity across instances based on relative information when absolute information (the constituent pitches) varies. Neuroscientific data suggest that the brains of absolute pitch possessors process pitch differently than do other brains. With respect to anatomy, musicians with absolute pitch have an asymmetry between the left and right hemispheres that prioritizes the left hemisphere. This asymmetry is found in an auditory association area known as the planum temporale, positioned towards the rear (posterior) of the temporal lobe (Schlaug, Jäncke, Huang, & Steinmetz, 1995). This finding is consistent with other research suggesting that the *lateralization* (that is, the degree to which a neural function is exclusive to the left or right hemisphere) of musical pitch varies with training.

We may infer from lesion/function deficit studies that the processing of musical pitch is predominantly a right-hemisphere activity. Since we know that pure pitch identification is a function of the primary auditory cortex, it is not surprising that damage to this region results in loss of simple pitch identification. However, whereas for most people musical pitch is processed predominantly in the right hemisphere, with the left dominant for language, trained musicians can show left dominance also for music (Bever & Chiarello, 1974). Thus the left hemisphere may be dominant for auditory processing that involves the application of internalized categories, such as absolute pitch. Yet another, more recent proposal is that the right hemisphere is optimized for analyzing the spectrum of a sound (which yields its pitch, see chapter 2), whereas the left hemisphere is optimized for perceiving rapid temporal fluctuations (common in speech; Zatorre, 2003).

Key

One of the most striking characteristics of musical pitch processing is the tendency to categorize pitch with respect to the surrounding tonal context (Krumhansl, 1990). In fact, there seem to be specific regions that serve the recognition of the sequences of notes that make up a scale. A recent study using fMRI attempted to identify brain regions responsible for detecting when a single tone 'pops out' based on its divergence from a tonal context (Janata, Birk, van Horn, Leman, Tillmann, & Bharucha, 2002). The researchers simply inserted tones that did not belong to the key, such as an F♯ in the key of C major. The study found consistent activation in the superior temporal gyrus (see Figure 4.2) – an area of general importance for pitch perception (Peretz & Zatorre, 2005), more so for the right than left hemisphere. However, the main focus of the authors was on what part of the brain 'tracks' changes in tonality, which they emphasized as being the rostromedial prefrontal cortex – an area just behind the center of one's forehead.

Other research concerning the way in which tonality makes certain pitches 'stand out' from the rest has applied the ERP technique. This research has revealed changes in electrical responses when a pitch does not match the overarching scale context. Interestingly, more recent work has shown that the brain responds to out-of-tune notes (which do not require a sense of scale, but only require an understanding of basic pitch categories) more quickly than unexpected notes that are appropriately tuned (Brattico, Tervaniemi, Näätänen, & Peretz, 2006). This finding suggests that the interpretation of pitch within a musical scale may constitute a higher-level cognitive task relative to detecting a mistuned note.

Sequential relationships

It is possible that particular neural processes are reserved for perceiving and/or remembering a pitch sequence as an integrated whole and for registering deviations from this. Much of the case study research reported by Isabelle Peretz and colleagues addresses this issue. In the paradigms used by this group, participants are typically presented with melodic sequences in pairs, in which the second sequence is identical to the first with the exception of one note, which may be altered in pitch or duration. Findings from this research suggest that damage to the right temporal lobe selectively disrupts the ability to detect deviations in melodic contour (Liégois-Chauvel, Peretz, Babaï, Laguitton, & Chauval, 1998; Peretz, 1990). Similarly, imaging techniques suggest a separation in the regions where melodies are recognized by overall pitch contour and where they are recognized by sequences of specific interval relations alone. The former tends to be dependent on a right-hemisphere activity, while the latter seems to involve regions in both hemispheres (Peretz & Zatorre, 2005). Peretz (1990) interpreted this finding as supportive of the view that the right hemisphere is involved in more 'holistic' processing whereas the left hemisphere is specialized for finer details.

Integration

As we opened the chapter with a case study, we conclude our discussion in a similar fashion. The case of Rachael Y., described by neurologist Oliver Sacks (2007), serves as a poignant example of the importance of our ability to integrate rich musical works into Gestalt wholes, an ability most of us take for granted. Rachael Y. was an accomplished middle-aged composer and performer when she suffered severe injuries to the head and spine following a serious car accident. (The locations of the injuries were not specified in Sacks's account.) Upon recovering from a coma that lasted several weeks, she found most abilities such as speech to be intact, but noticed a change in her perception of music. She described her experience of the first piece she heard following her recovery from the coma (Beethoven's opus 131):

> When the music arrived, I listened to the first solo phrase of the first violin again and again, not really being able to connect its two parts. When I listened to the rest of the movement, I heard four separate voices, four thin, sharp laser beams, beaming to four different directions. Today, almost eight years after the accident, I still hear the four laser beams equally . . . and when I listen to an orchestra I hear twenty intense laser voices. It is extremely difficult to integrate all these different voices into some entity that makes sense. (p. 113)

As we have discussed, melody is not perceived in isolation. It does not only require the perception of a horizontal sequence, but the complex interweaving of all the parts including the vertical axis (harmony). Having lost the basic ability to integrate the many rich parts of music into a coherent whole, Rachael Y. experienced music as quite unpleasant and chaotic, requiring 'a great cognitive effort to hold the strands together' (p. 116).

Coda

In many ways, the perception of melody is central to our experience of music. GL's case illustrates the dramatic loss of ability to recognize the storehouse of tunes we accumulate throughout our lives. Without the ability to conceptualize tone sequences as gestalts, and to hear the tones as part of a coherent tonal system, music becomes meaningless. At the same time, you may have noticed that the term 'melody' in this discussion has almost exclusively referred to pitch. It is common for those in music cognition to use the term melody in a way that excludes rhythm. The assumption here is that rhythmic relationships contribute independently to musical experience. But do they? We consider this, and other issues related to rhythm, in chapter 6.

Notes

1 See also Satoh, Takeda, and Kuzuhara (2007).

2 This anecdote exists in various versions and is attributed to many different composers, including Mozart. The authors were unable to find a reliable print source and assume the often-told tale to be apocryphal!

3 DeWitt and Samuel's (1990) extensive paper reports a series of five experiments (four using melody lines and one employing chords). The findings of the five studies taken as a whole are more complex than reported here, and the interested reader is referred to the original paper for more detail.

4 Not all participants heard exactly the same percept. Several examples of how the pattern is perceived are shown in the liner notes to Deutsch's (1995) CD. Some participants hear only a single pitch (same pitch 'heard' in both ears) so that two tones are not even perceived. Despite some variation in responses, none of the reports correspond to the actual notes played in each ear.

5 As might be expected, responses to probe tones are highly related to the key suggested by the preceding context, and responses are more distinct for keys that are more distant according to the circle of fifths (see Figure 5.4). For instance, the response to an F# probe tone is much higher (indicating greater stability) after hearing a context in the key of F# than the key of C. Along with Mark Schmuckler, Krumhansl devised a mathematical algorithm that can be used to determine a listener's perception of key by using responses to a series of probe tones (described in Krumhansl, 1990, Chapter 4). This algorithm has been highly reliable and is used often in the literature (see e.g., Temperley, 2001).