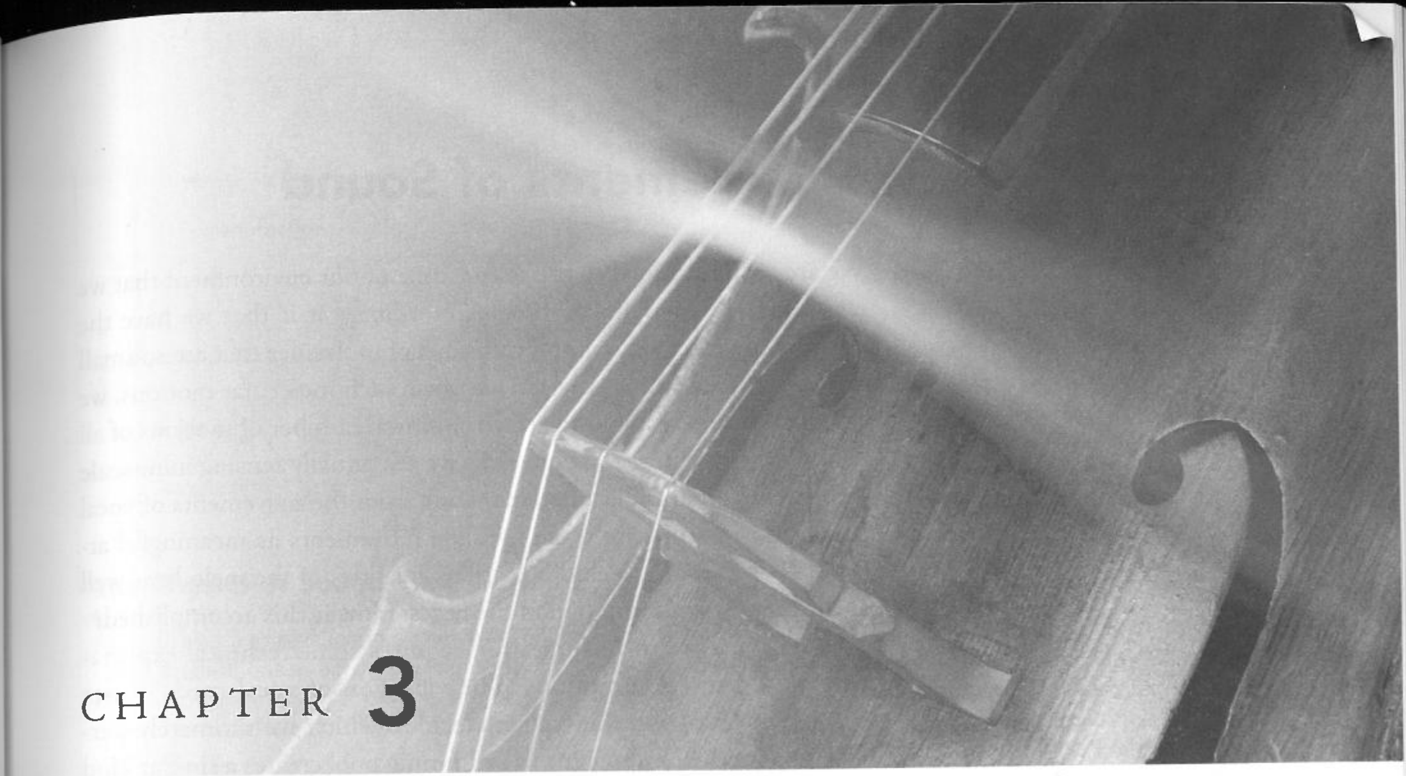


CHAPTER 3

Musical Building Blocks

LEARNING OUTCOMES

By the end of this chapter you should be able to:

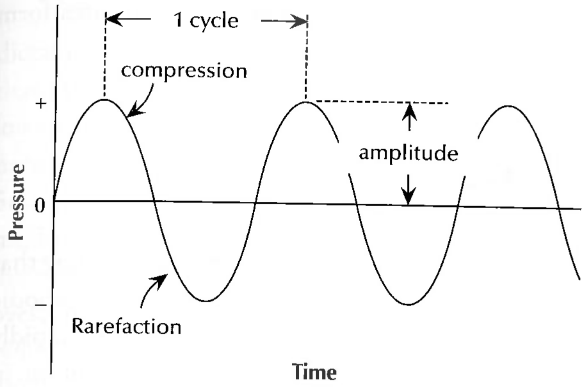
1. Describe how vibrating objects give rise to sound, and how sound is detected by the auditory system.
2. Explain the physical basis of periodic and nonperiodic sounds.
3. Discuss the connection between music and the acoustic structure of sound.
4. Identify the psychoacoustic effects of combining musical tones and outline the implications of these effects for music.
5. Contrast and evaluate psychological models of pitch.
6. Describe properties of sound that influence our perception of timbre.

The Elements of Sound

We are so accustomed to the soundscape of our environment that we rarely stop to contemplate how extraordinary it is that we have the capacity to detect actions and collisions of molecules that are so small as to be invisible. Not only can we sense such molecular motions, we can identify and respond to an astronomical number of motions of all types. When we hear someone sing, we are actually sensing minuscule movements of air molecules emanating from the movements of vocal cords, and interpreting these invisible movements as meaningful attributes of the song, including various qualities of the melody as well as the linguistic information in the lyrics. How is this accomplished?

The answer to this question requires some technical explanations and a few assumptions. The vibration of vocal cords causes movement in surrounding air molecules, which we ultimately perceive as sound. (Jumping into a swimming pool creates a similar kind of wave, but on a larger scale.) This movement of air molecules emanates outward from a vibrating source and eventually collides with the *eardrum*. The movements detected by the ear are infinitesimal but powerful, for they are ultimately responsible for our auditory world, including our most meaningful music experiences.

When the eardrum responds to the movements of air molecules, a complex chain of events is initiated that implicates physiological and neural activity at many levels of processing. Vibrations of the eardrum lead to movement of the *basilar membrane*, a resonant structure that resides within the *cochlea* of the inner ear. The basilar membrane is not unlike the string of a guitar, in that it varies in width and stiffness and vibrates in conjunction with sound. Unlike any musical instrument, though, it is surrounded by liquid and vibrates within that medium. The actual motion of the basilar membrane is described as a *traveling wave* or progressive wave. Other examples of traveling waves include ocean waves, the movement that occurs when a whip is cracked, and audience waves at sporting events. Each point along the basilar membrane has a *characteristic frequency* (CF), which is the frequency for which it is most resonant. Thus, although any sound will cause a traveling wave along the basilar membrane, the maximum amount of movement in that wave will occur at points that depend on the frequencies involved in the sound. The basilar membrane is the base for the sensory receptors of hearing: the *hair cells*. Movement



The Physics of Sound: Sine Waves

Air has mass and elasticity, so any source of energy will cause fluctuations in air pressure in the form of waves. At a given point along the wave, the air molecules will either be bunched up (*compression*) or spread apart (*rarefaction*).

Shown here is a graphic illustration of the alternating pattern of compression and rarefaction of the simplest type of sound wave, called a *sinusoidal* or *sine* wave. The x-axis indicates the unfolding of time, and the y-axis indicates the increase (+) and decrease (–) in sound pressure taking place.

A *period* is the amount of time that it takes for a complete cycle of the waveform. In the diagram, it is demarcated as the "1 cycle" area from one compression peak to the next.

Amplitude refers to the intensity of the fluctuation in air pressure, illustrated in the diagram by the arrow indicating the maximum point of compression. It is most commonly measured in decibels (dB), and its psychological correlate is loudness. A louder sound is produced by waves with greater amplitude.

Repetition rate is measured in hertz (Hz), the number of cycles completed in one second. A period equals the reciprocal of the repetition rate. If the repetition rate is 100 Hz (100 cycles/second), then the period is 1/100 second.

Two sine waves equal in period will interact according to their *phase*. If they are 180° out of phase (the compression peak of one wave coincides with the rarefaction peak of the other), they will tend to cancel each other out. If they are 180° out of phase and of equal amplitude, silence will result.

Source: Reproduced from Schiffman, H. R. (2000). *Sensation and perception: An integrated approach.* New York: Wiley, p. 316. Used by permission.

of the basilar membrane causes movement of hair cells, leading to electrical signals that are relayed from the auditory nerve to the auditory brain stem and eventually the cortex. This extraordinary mechanism for detecting the movement of air molecules allows human listeners to distinguish roughly 1,500 frequencies.

Why do certain types of movements give rise to rich experiences of music, and others give rise to experiences of speech, traffic noise, or laughter? The answer lies in an examination of the sensory attributes of sound. When sound stimulates the auditory system, it is subject to a kind of physiological interpretation or analysis, whereby vibrating molecules are converted into various forms of neural activity that represent the sound in terms of a range of sensory attributes, a process called *sensory transduction*. These sensory attributes form the building blocks of music.

Periodic Motion

A lot of the movement that occurs in our environment, including that of vocal cords when we talk or sing, is called *periodic motion*. Periodic motion occurs when an object such as a plucked guitar string rapidly moves back and forth in a repetitive manner. This type of motion, in turn, causes a periodic sound wave. When a periodic sound wave reaches the eardrum, it gives rise to a distinctive quality that is highly relevant to music: the sensation of *pitch*.

Any naturally occurring sound, such as that of a footstep, a sneeze, breaking glass, or a note played on the flute or piano, can be described as a combination of many simple periodic waves (i.e., sine waves) or *partials*, each with its own frequency of vibration, amplitude, and phase. *Fourier analysis* is the technique that allows us to analyze a complex sound into its sine-wave components. The various sine-wave components of a complex sound are collectively referred to as the *spectrum* of frequency components, or simply, the *sound spectrum*.

We do not normally perceive the individual partials of a complex sound because cognitive mechanisms operate to fuse them together, leading us to experience a unitary sound. According to *Ohm's acoustical law*, however, under certain listening conditions we do have a limited ability to hear some of the individual partials of a complex sound. For example, when a single piano note is played, under certain conditions it is actually possible to hear some of the many pitches (i.e., partials) that make up the note.

Sounds with a discernible pitch all have a *periodic waveform*, in that the waveform continuously repeats itself over time. For example, playing an individual note on a flute, piano, violin, or any other "pitched" instrument will generate a periodic waveform. The *period* is the time taken for one complete cycle of the waveform; it is the reciprocal of the repetition rate. The *repetition rate* usually determines the perceived pitch (exceptions include circular tones, discussed later) and is measured in cycles per second, or hertz (1 Hz = 1 cycle/sec). In other words, when a periodic waveform impinges on the auditory system, it is ultimately converted into a sensation of pitch. In general, if the repetition rate of a periodic sound is increased, the perceived pitch increases.

The partials of any periodic waveform tend to occur in a predictable pattern of frequencies known as the *harmonic series*. Partials that fall along the harmonic series are also called *harmonics*. If the lowest frequency component of a periodic sound is n, then the other harmonics—called overtones—are members of the set $2n$, $3n$, $4n$, $5n$, and so on. That is, each overtone has a frequency that is an integer multiple of the lowest or fundamental frequency of the complex. An important property of the harmonic series is that additional harmonic overtones do not alter the overall repetition rate of the waveform (which is determined by the fundamental frequency) and therefore do not change the perceived pitch of the complex.

The Discovery of Music Within Sound

The mathematical precision of the harmonic series is actually an idealized description of complex periodic sounds, and naturally produced sounds typically include inharmonic components. Moreover, depending on the source of the sound (e.g., voice, piano, violin) and the manner and context in which that sound is produced (intensity and room acoustics), different harmonics will be more or less prominent. Nevertheless, the frequency relationships that are represented in periodic sounds are remarkably musical. Helmholtz (1863/1954) observed that several aspects of music, notably scales and harmony, have compelling parallels in the acoustical structure of sound. He noted that music from several cultures involves important scale notes that map onto the harmonics of complex tones. In the major scale, the fifth scale degree (sol) is equivalent (i.e., in note name, or *tone chroma*) to the third harmonic of a complex periodic tone built on the first note of the scale (doh), and the third scale degree (mi) is equivalent to the fifth harmonic.

Remarkably, the most important musical intervals that are used in Western music can be found within the harmonic content of a single note. The first and second harmonics of a complex tone are separated by an *octave*, the second and third harmonics are separated by a *perfect fifth*, the third and fourth harmonics are separated by a *perfect fourth*, the fourth and fifth harmonics are separated by a *major third*, and the fifth and sixth harmonics are separated by a *minor third*. How could such a mapping occur? In spite of Ohm's acoustical law, it is not easy or typical to hear the individual frequencies contained within an individual note.

Sound Example 3.1 is an illustration of the harmonic series using a guitar string. The fundamental pitch of the open plucked fifth string is an A at 110 Hz. If the string is lightly fretted at the 12th fret and plucked, the second harmonic, an A an octave above and twice the frequency of the fundamental (220 Hz), is heard. The third harmonic, an E a perfect fifth above the second harmonic and three times the frequency of the fundamental (330 Hz), is found at the seventh fret. The fourth, fifth, and sixth harmonics can be found at the fifth, fourth, and third frets, respectively.

When fretting a harmonic, the physical division of the string reflects the frequency ratio of harmonic to fundamental. For example, the second harmonic forms a ratio of 2:1 and is found at the exact midpoint of the length of the string (the 12th fret). The third harmonic forms a 3:1 ratio with the fundamental and is fretted at a point that is exactly a third of the length of the string (the seventh fret).

Source: © 2008, Richard Yates.

Rather, the auditory system seems to fuse all of them together, giving rise to a single and unified perceptual experience of that note.

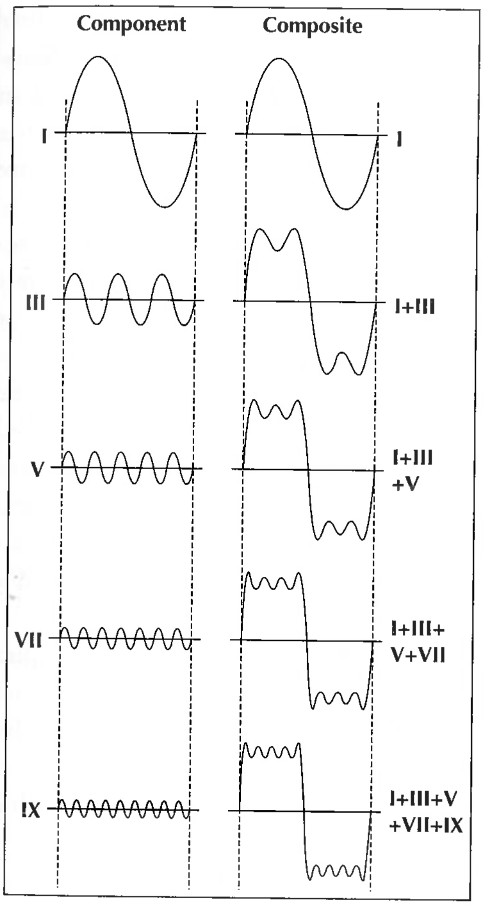
In spite of almost 150 years of scholarship on musical acoustics, it remains an intriguing mystery as to how or why the harmonic spectra contained within individual tones are mirrored in the many pitch relationships that can be found in our melodies and harmonies. The predominance of musical relationships that seem to echo the very frequency relationships contained within individual complex tones implies that, at some level, the brain must register frequency relationships within individual sounds. At a conscious level, however, we do not usually perceive frequency relationships within individual tones. If we play a single key on a piano, we hear a single note, not an array of frequency relationships. Rather than generating individual pitch sensations, the spectral content of individual tones is fused and contributes to another attribute of sound that seems to have little to do with musical relationships: the quality of timbre. That is, the spectral content of a note is perceived as having particular sound qualities, tone colors, or timbres.

Our perceptions of sound, it seems, are deceptively oblique, and it is only through careful investigation that we have any chance of fully understanding the connection between the physics of sound and the nature of music experience. Pythagoras (6th century B.C.) laid much of the groundwork for this quest with his discoveries on the nature of sound, but it is only in the past 100 years, following many advances in the study of human physiology, that we have been able to move significantly beyond these insights to gain an understanding of the perceptual and cognitive basis of music.

Sensory Consonance and Dissonance

The spectrum of harmonic overtones not only influences our perception of timbre; it also affects the degree of sensory consonance and dissonance that we perceive when multiple tones are presented simultaneously. Long-term knowledge of music also affects perceptions of consonance and dissonance, and these qualities of music are referred to as *musical consonance* and *musical dissonance*. Sensory and musical forms of consonance and dissonance are often overlapping, but according to the most prominent theories they result from distinctive mechanisms.

Whereas sensory consonance and dissonance are thought to result from psychoacoustic factors and arise independently of learning and enculturation, musical consonance and dissonance are thought to be dependent on musical experience and can change throughout the life span.



The Physics of Sound: Fourier Analysis

Any complex sound can be broken down into a set of component sine waves called partials using a mathematical procedure called Fourier analysis, named after its discoverer, the French mathematician Jean Baptiste Fourier (1768–1830). The process can also be reversed: Complex sounds can be artificially constructed from a set of sine waves (partials) by a process termed *Fourier synthesis*.

This figure illustrates a synthesized periodic waveform called a square wave, produced by adding only the odd-numbered partials in the harmonic series to a fundamental frequency. For example, if a fundamental frequency (1) is 100 Hz, (3) is 300 Hz, (5) is 500 Hz, and so on. The amplitudes of the partials are also proportional to the fundamental frequency. That is, (3) is 1/3 the amplitude of (1), and (5) is 1/5 the amplitude of (1). The left column lists each component harmonic separately, and the right column illustrates the effect on the composite wave as each harmonic is added.

Source: Reproduced from Boring, E., Langfeld, H., and Weld, H. (1948) *Foundations of Psychology*. New York: Wiley, p. 316. Used by permission.

Sound Example 3.2

An illustration of the critical band. Two sine tones are played at the same frequency; one remains constant and the other slowly increases in frequency. As the difference between the tones passes through the critical band (one third of an octave), an interference effect known as *beats* occurs. This is followed by a roughness or dissonance that slowly dissipates. The rising tone stabilizes at a simple integer ratio of 3:2 (perfect fifth) from the constant tone, and the roughness disappears.

Intervals that give rise to sensory consonance are made up of tones with fundamental frequencies that stand in small integer ratios. The most consonant interval, the octave, consists of pitches with fundamental frequencies that are related to each other by a ratio of 2:1. Another highly consonant interval, the perfect fifth, consists of pitches with fundamental frequencies that are related to one another by a ratio of 3:2. In contrast, intervals that give rise to sensory dissonance comprise tones with fundamental frequencies that stand in more complex ratios. For example, the fundamental frequencies associated with the dissonant "tritone" interval (six semitones) are related to each other by a ratio of 45:32.

Tone combinations with fundamental frequencies that are related by small integer ratios, such as the octave (2:1) and the perfect fifth (3:2), have many overtones that either align with one another or are separated in frequency by a distance that is greater than a *critical band*. The concept of a critical band originated from psychoacoustic research and refers to the range of frequencies within which so-called *sensory interactions* occur. Across the pitches that are typically used in music, the critical band corresponds to about one-third of an octave. Consonant intervals do not give rise to many such sensory interactions, and the basilar membrane can largely resolve the many frequency components associated with such pitch combinations.

In contrast, combinations of pitches that are related by complex ratios contain many overtones that are not aligned with one another but that are too close in frequency for the basilar membrane to resolve optimally. Such combinations contain a lot of harmonic frequencies that fall within a critical band, resulting in sensory interactions that include rapid amplitude fluctuations. These amplitude fluctuations seem to give rise to a sensation of roughness and beating, and the perception of dissonance. Dissonant intervals give rise to a lot of sensory interactions, and the basilar membrane is not able to resolve all of the frequency components associated with these pitch combinations.

Because the frequency components of dissonant intervals are poorly resolved by the basilar membrane, dissonant intervals are also discriminated poorly by infants, and even by adults. Just as it is difficult to recognize a face in a blurry photograph, so too most people are not very good at judging dissonant sounds. These marked differences in responses by peripheral auditory mechanisms to consonance and dissonance, including the sensation of roughness, are a simple means for humans and various nonhuman species to discriminate consonance and dissonance. Interestingly, however, *preferences* for consonance appear to be uniquely human (McDermott & Hauser, 2005).

The prevailing view is that consonance is equivalent to a "smooth sound," which occurs when there is an absence of any sensation of roughness and beating. However, there may be other influences on our perceptions of consonance. One view is that we perceive combinations of tones to be consonant if the combined set of frequency components can potentially arise from a single complex tone. That is, if two tones played at the same time sound like a single tone then the combination is perceived to be consonant. Why would two simultaneous tones sound like a single tone? As described, frequency components that stand in a harmonic relationship tend to *fuse* into a single perceptual unit, regardless of whether those frequency components arise from the same source.

Take the octave interval. When notes separated by an octave are combined, the frequency components of the higher-pitch tone coincide exactly with all of the even-numbered frequency components of the lower-pitch tone. The result of this combination is a new complex tone that, like the two component tones, aligns with the harmonic series and sounds like a single tone. *Harmonicity* is the degree to which any set of frequencies aligns with the harmonic series (f, 2f, 3f, 4f, 5f, etc.). The octave interval has high harmonicity and sounds consonant. Other tone combinations, such as the augmented fourth, have low harmonicity and sound dissonant.

Does consonance arise from an absence of roughness and beating, or from a high level of harmonicity? The question is difficult to answer because tone combinations with high harmonicity virtually always have a low level of roughness. In other words, the two factors are *confounded*. To disentangle these two possible influences on consonance, McDermott, Lehr, and Oxenham (2010) collected preference ratings for tone combinations that varied in the amount of beating but *not* harmonicity. Because beating arises in the peripheral auditory system, there is less of a sensation of beating when pure tones in a dissonant relationship are presented to *different* ears than when they are presented to the same ears, but the degree of harmonicity remains the same in both circumstances. By comparing preference ratings for pure tones presented to the same ears or to different ears, they were able to isolate the effect of beating on preference ratings. They next obtained preference ratings for sounds that varied in harmonicity but not beating. These sounds consisted of tones with harmonic and inharmonic components that could not physically generate beats, because the components all differed from each other by more than a critical band. Finally, they obtained preference ratings for numerous tone combinations that varied in consonance. Although people generally like consonant tone combinations, people vary in how tolerant they are to dissonance. The question was whether individuals who showed the strongest preferences for consonant tone combinations also showed similar preferences for either harmonicity or the absence of beating (or both).

Interestingly, a dislike of beating did not predict preference ratings for consonant tone combinations, whereas a strong preference for harmonicity seemed to align well with a strong preference for consonance. These results support the idea that consonance is determined by harmonicity rather than beating and roughness. However, more research is needed before the nature of consonance is fully understood. Other research, for example, suggests that consonance is determined by neither harmonicity nor beating, but by enculturation and learning (see McLachlan, Marco, Light, & Wilson, 2013; McLachlan & Wilson, 2010). In other words, consonant sounds might just be familiar sounds.

Sound Example 3.3

The sa and pa in the Indian classical tradition are equivalent to do and sol in Western nomenclature. They are sounded continuously (drones) on the *tambura* during a performance.

Source: Martin Quibell. Made available under Creative Commons Sampling-Plus License 1.0.

Tuning Systems

If sensitivity to sensory consonance is a basic property of the auditory system, one might ask how this property affects scale structures and tuning systems. In most scales from around the world, consonant intervals (e.g., octaves, perfect fifths and fourths) are structurally important. For example, tones separated by octaves are considered to be similar in virtually all musical cultures. In North Indian scales, tones separated by a fifth (the sa and pa) are structurally important and are typically sounded continuously throughout a piece. The most common pentatonic scale (exemplified by the black notes on the piano), which is found in Chinese and Celtic music, can be formed by choosing any pitch as an arbitrary starting tone, adding a second tone a fifth higher, another tone a fifth higher than the second tone, and so on, until a collection of five pitches is obtained. The scale is formed by octave-transposing the collection of tones so that they fall within a single octave.

Sound Example 3.4

A major third interval is presented melodically in three tunings: equal temperament (most commonly used today), Pythagorean, and just intonations.

In Western music, two similarly natural tuning systems for the chromatic scale have been used historically (for a review, see Burns, 1999). One, called Pythagorean tuning, extends the pentatonic scale already described with additional tones that continue the cycle of fifths. In the other, called just intonation, the scale is formed by tuning notes so that their fundamental frequencies form small integer ratios with the fundamental frequency of the first note of the scale (doh). Both of these scales limit the possibility of transpositions between keys, because some instances of particular intervals (e.g., the perfect fifth between C and G) are tuned differently than other instances (e.g., C# and G#).

Equal temperament represents a compromise solution. It guarantees that all intervals (e.g., all perfect fifths, or all major thirds) are tuned identically, and that important intervals do not deviate greatly from small integer frequency ratios: Fifths and fourths deviate from exact small integer ratios by 2% of a semitone; major and minor thirds are slightly more mistuned. These minor deviations, although discriminable in some cases, are no greater than the typical tuning deviations observed in the performances of singers or stringed-instrument players. Moreover, minor departures from exact small integer ratios have little effect on the perceived consonance of these intervals, which may explain why equal temperament has endured for many years.

Not all musical scales are based on consonant intervals. As two examples, the slendro and pelog scales, used in Gamelan music of Bali and Java, are strikingly different from Western scales. A possible reason is that Gamelan music tends to feature instruments with inharmonic spectra, such as xylophones, drums, and gongs. The timbre of instruments that predominate in a musical culture, in turn, influences how those instruments are tuned and, hence, the scales associated with the music. In traditions that emphasize instruments with inharmonic spectra, scales are very different from Western

J. S. Bach and Equal Temperament

The pioneering endorsement of equal temperament on the part of J. S. Bach (1685–1750) was initiated to "expand the harmonic universe." The problem with both just and Pythagorean tuning systems is that, of the 24 major and minor scales commonly used by Western musicians at the time, only a limited number remained viable. For some scales, certain intervals would sound noticeably out of tune. As a consequence, some musicians would be entirely confined to a few scales; others would have to retune their instrument to be able to play using other scales. Equal temperament, the division of the octave into 12 equal parts, allowed musicians to borrow from any scale at any time without having to retune. Bach's *Well-Tempered Clavier* was a set of 48 keyboard pieces, two in each major and minor key, designed to illustrate the virtues of equal temperament.

diatonic major and minor scales. This difference arises because scales are designed not just for the purpose of creating melodies, but also to support simultaneous combinations of notes.

For example, a gong that has a spectrum with a prominent tritone will tend to be tuned to a scale that features a tritone. That way, a gong tuned to the first scale note (which itself features a tritone frequency component) will combine well with a second gong tuned to the tritone. Both gongs feature a tritone in their frequency spectrum, and this shared component helps to fuse the notes together. To illustrate, imagine an artificial tone (created using a synthesizer) that consists of only two inharmonic partials at frequencies f and $\sqrt{2}$f (where the distance of $\sqrt{2}$f corresponds to a tritone, or six semitones in an equally tempered chromatic scale). The spectrum of this artificial tone is inharmonic (the partials are not compatible with the pattern f, 2f, 3f, 4f . . .) and therefore it does not give rise to a very clear pitch sensation. Instead, it sounds a bit like a chime. As shown in Figure 3.1, if this *tritone chime* is combined with another tritone chime at progressively divergent pitch distances, the degree of roughness and beating is minimized at zero semitones, six semitones (the tritone), and 12 semitones (the octave). In theory, such an instrument should work best in a scale that includes these notes.

Arguing along these lines, Sethares (2005) argued that tuning systems and scales co-evolve with musical instruments. Instruments that are played together are tuned in a way that supports their combination. Once established, a tuning system can accommodate the development of new instruments that have spectral properties consistent with the tuning system. In other words, the tuning system that works best for one instrument also constrains new instruments that are developed. This

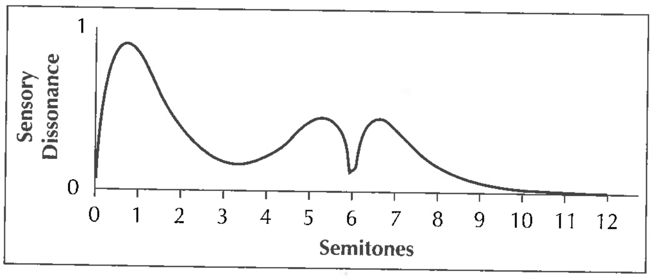


Figure 3.1

Illustration of sensory dissonance when two tritone chimes are combined at progressively divergent pitch distances.

process of co-evolution explains why Gamelan instruments such as the boning and saron are rarely used with the scales of Western music, but they work well with the slendro and pelog scales.

Sensitivity to Pitch

People are remarkably sensitive to pitch, but this sensitivity relates primarily to pitch relations. Sensitivity to pitch relations, often called *relative pitch*, refers to the ability to produce, recognize, or identify pitch relations, such as those that define a musical interval or a melody. To illustrate, "Happy Birthday" can be sung in a high or a low voice, or performed on a tuba or a piccolo. As long as the pitch relations conform to those of the melody, it is recognizable. People are much less sensitive to isolated pitches. Anyone can appreciate the quality of an isolated timbre (e.g., flute, cello, oboe, trumpet), but an isolated pitch leaves us cold. Only in exceptional cases do individuals form rich associations with isolated pitches.

In a case study reported by Luria, Tsvetkova, and Futer (1965), a person referred to as S. formed detailed visual images on hearing individual pitches. When presented with a 64 dB tone pitched at 250 Hz, he saw a velvet cord with fibers jutting out on all sides. When presented with a 100 dB tone pitched at 500 Hz, he saw a streak of lightning that split the sky in two. Other cases of synesthetic experiences of pitch have also been reported, but the vast majority of people have simpler experiences. We might be able to assert that an isolated pitch sounds high or low, but the most interesting psychological qualities of pitch emerge from the relationships that are formed when we hear multiple pitches.

Pitch is implicated at multiple levels of musical structure, forming the basis for melody, harmony, and key. As we listen to music, we experience pitch on all of these levels, giving us a multidimensional

Musical Schemata

What schemata do we use to place music into categories? Some categories such as rap and opera seem obvious. Others have blurry boundaries. What is the difference between pop and rock, or rock and heavy metal, or heavy metal and punk?

A recent music marketing trend may provide some empirical data on this issue. In an attempt to appeal to as many radio station formats as possible, record companies are releasing multiple "mixes" of songs. Shania Twain's entire CD *Up!* was released in two mixes: a country version and a pop version. Examining the differences between mixes might provide some clues as to what constitutes our mental representations of these various categories.

understanding of structure. Our understanding of music is stored in the brain as a mental representation. Mental representations of music are partly built up from experience. Through passive or active exposure, listeners gradually internalize regularities in the music of their own culture, forming long-term knowledge schemata into which novel music stimuli are assimilated. Nonetheless, mental representations of pitch may also be constrained by processing limits and biases. For example, the limits of working memory may constrain our ability to encode unfamiliar melodies with accurate detail. Because of such limits, mental representations of unfamiliar melodies are sketchy, often containing little more than pitch contour information.

Models of Pitch Perception

Pitch height is the most basic dimension along which pitches are perceived to vary; it refers to the continuum that extends from low to high pitches (which corresponds to a logarithmic function of frequency, or cycles per second). The psychological relevance of this continuum is evident in similarity judgments for pairs of pitches: Tones closer in pitch are considered more similar than tones separated by greater pitch distance. The pitch-height continuum is also evident in neural activity in the cochlea. High frequencies stimulate the basal portions of the basilar membrane, low frequencies stimulate the apical portions, and intermediate frequencies affect intermediate portions.

In constructing a model of pitch perception, one may start with the basic dimension of pitch height, and then consider additional dimensions along which pitch seems to vary. Most models of pitch perception consider the special status of the octave among intervals. As noted, tones with fundamental frequencies separated by an octave are perceived to be similar in virtually all cultures. Pitch chroma refers to the quality of pitch that is independent of the octave register in which it occurs. In Western music, tones separated by an octave are given the same name (e.g., A, B, C), implicating their equivalence in some sense. Listeners with musical training are especially sensitive to the similarity between tones that are separated by an octave (Allen, 1967; Kallman, 1982).

If pitch chroma and pitch height are basic dimensions of pitch perception, we should expect to find similar evidence among naïve listeners. When musically untrained adults or children judge the similarity between pure tones, however, they tend to focus exclusively on the dimension of pitch height (Allen, 1967; Kallman, 1982; Sergeant, 1983). For example, C4 and C#4 (notes separated by a semitone) are perceived to be highly similar, but C4 and C5 (notes separated by an octave) are perceived to be no more similar than C4 and B4 (notes separated by a major seventh). These findings do not rule out the possibility that sensitivity to the chroma dimension could be uncovered with tasks that measure implicit rather than explicit knowledge of musical associations. However, the findings make it clear that music educators should not assume naïve listeners will have an explicit understanding of octave equivalence.

An early psychological model of pitch perception incorporated the two dimensions of pitch height and pitch chroma. As shown in Figure 3.2, these two dimensions are depicted as orthogonal dimensions of a geometrically regular helix—a monotonic dimension of pitch height and a circular dimension of pitch chroma. Shepard (1964) reported evidence that these dimensions are psychologically relevant and orthogonal. He created tones with well-defined chroma but ambiguous height. Such circular tones were constructed by combining 10 pure-tone components spaced at octave intervals, and imposing a fixed amplitude envelope over the frequency range such that components at the low and high ends of the range approach hearing threshold.

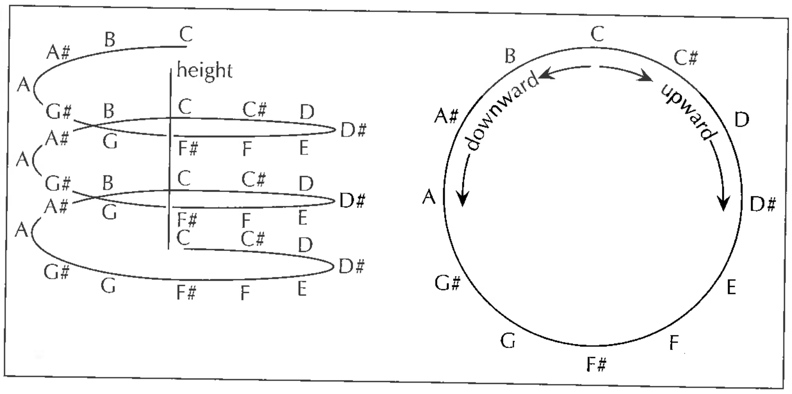
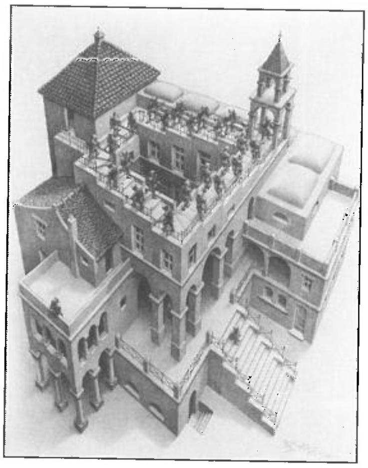


Figure 3.2

An early model of pitch perception. Two psychological dimensions are represented in the model: height and chroma.

Sound Example 3.5 is a scale comprising Shepard tones. Most listeners perceive the scale as ascending because it implies smaller pitch distances from tone to tone, even though pitch height has been rendered ambiguous. Because pitch ascension is illusory, this excerpt could conceivably be looped to create the illusion of a scale endlessly rising in pitch. A visual analogy can be made with M. C. Escher's "Ascending and Descending" (below). Because the relative height of any one stair in the staircase is ambiguous, an illusion of an endlessly ascending (or descending) staircase is created.

Source: M. C. Escher's "Ascending and Descending" © 2008 The M. C. Escher Company-Holland. All rights reserved. www.mcescher.com



Although the overall pitch height of any circular tone is somewhat indeterminate, listeners experience certain circular tones as being higher or lower than others. In particular, listeners tend to perceive the relative height of these tones so as to maximize their pitch proximity. For example, when presented the circular tones C followed by D, one might perceive the second tone as higher or lower than the first tone (C up to D, or C down to D). Both interpretations are possible because their individual pitch heights are ambiguous. Nonetheless, listeners typically perceive the second tone (D) as higher than the first tone (C), because this interpretation implicates a pitch distance of only two semitones. Listeners almost never perceive the second tone as lower than the first tone, because that interpretation would imply a pitch distance of 10 semitones.

Shepard created fascinating patterns of circular tones in which chroma varied continuously around a chroma circle (see Figure 3.2). For ascending patterns, he shifted all of the octave spaced components up in frequency (adding new components at the low end of the amplitude envelope) until the complex returned to the initial configuration. When chroma was shifted continuously in a clockwise direction around the chroma circle (C, C#, D, D#, etc.), the pattern was perceived as ascending endlessly. When chroma was shifted continuously in a counterclockwise direction (C, B, A#, A, etc.), the pattern was perceived as descending endlessly. That is, listeners perceived changes in pitch chroma but not in overall pitch height. Most important, these effects were perceived by musically trained and untrained listeners alike. By making pitch height indeterminate, Shepard demonstrated that pitch chroma has perceptual significance, even for untrained listeners. Thus, although untrained listeners may lack explicit knowledge of the similarity between tones with the same chroma, they nonetheless demonstrate sensitivity to chroma.

If pitch class and pitch height are orthogonal, then the relative height of circular tones that are directly opposite on the circle, such as C and F#, should be ambiguous or indeterminate. Deutsch (1999) addressed this prediction in a series of experiments. She reported that, for a given individual, certain circular tones (i.e., chromas) are reliably and consistently judged as being higher than the circular tone that is opposite on the circle, suggesting that the perceived height of a tone is systematically related to its pitch chroma. For example, some listeners perceive a successive presentation of two particular circular tones separated by a tritone (e.g., C and F#) as upward pitch movement, whereas others perceive the same two tones as downward pitch movement. In another investigation, Repp and Knoblich (2007) showed that performing a simple hand movement can affect whether sequential circular tones separated by a tritone are perceived as ascending or descending.

To date, no researcher has been able to construct a complex tone such that changes in pitch chroma do not generate corresponding changes in pitch height. For circular tones, perceived changes in pitch height are dependent on context, but they are never completely absent. In fact, some listeners report that sequences of circular tones give rise to *multiple* sensations of changing pitch height—some ascending and others simultaneously descending. That is, tones constructed to have an ambiguous pitch height may sometimes break apart into individual pitch sensations, but they will always generate one or more sensations of pitch height.

Shepard (1982, 1999) and others have proposed models of pitch perception with additional dimensions besides those based on height and chroma. These models account for affinities between tones separated by perfect fifths and fourths, which are said to be relevant for musically trained listeners tested with harmonically rich tones presented in musical contexts. Even more elaborate models have been proposed to account for affinities between tones separated by thirds. The establishment of a musical key also shapes pitch judgments. For example, in the context of the key of C major, C and G are more strongly associated with each other than E and B, even though both tone pairs represent a perfect fifth interval (Krumhansl, 1979).

The effects of musical context on judgments of pitch were examined in a series of investigations conducted by Krumhansl and her colleagues (see Krumhansl, 1990). They used the probe tone technique in a variety of musical contexts. The method involves presentation of a musical stimulus that clearly defines a musical key (e.g., scale, cadence, chord sequence) followed by a probe tone. Listeners rate (usually on a scale from 1 to 7) how well each probe tone in the chromatic scale fits with the established key.

When musically trained listeners are tested with this method, probe tone ratings are quite consistent across listeners. As we might expect, their ratings mirror predictions from music theory. When a major key is established, a probe tone corresponding to the tonic of the key has the highest rating, followed by the dominant and mediant, then the other tones in the key (diatonic or scale tones), and

Sound Examples 3.6.1 and 3.6.2 are demonstrations of the probe tone technique. A sequence of notes or chords is presented, followed by a single tone or "probe tone". The listener's task is to rate from 1 to 7 how well the final probe tone fits contextually, 1 being a poor fit and 7 being an excellent fit. Krumhansl and Kessler (1982) derived the tonal hierarchy from the responses of a sample of participants with a range of musical experience.

finally the nondiatonic (i.e., nonscale) tones. This tonal hierarchy is more easily uncovered with musically trained participants than untrained, illustrating the influence of learning. School-age children's implicit knowledge of the hierarchy improves dramatically from six to 11 years of age (Krumhansl & Keil, 1982), but even six-year-olds know that different tones vary in goodness once a musical context is established (Cuddy & Badertscher, 1987). Thus, although sensory dissonance is perceived by infants and musically untrained adults, judgments of pitch relations are also influenced by age and experience, suggesting that our perceptions of pitch and pitch relations result from a combination of innate and learned factors.

Sensitivity to the tonal hierarchy early in development can be explained—in part—by psychoacoustic influences (Schellenberg & Trehub, 1994). Probe tones with high stability (doh, mi, sol, fa) have the largest degree of sensory consonance with the established tonic. Alternatively, because the frequency of occurrence of tones in real pieces of music closely mirrors the tonal hierarchy, children may learn implicitly that tones heard more often in a musical piece are particularly stable. Indeed, we know that listeners are highly sensitive to pitch distributional information in music. Oram and Cuddy (1995) presented listeners with atonal sequences in which one tone occurred eight times, another tone occurred four times, and four other tones occurred once each. Following each sequence, listeners rated the extent to which various probe tones fit with the sequence in a musical sense. Ratings reflected the frequency with which each pitch occurred in the context. Thus, when listening to unfamiliar music, listeners readily construct a hierarchy of pitch importance using by attending to the frequency of occurrence of pitches.

Absolute Pitch

It is often asserted that listeners are largely insensitive to *absolute pitch*. Absolute pitch refers to the ability to produce, recognize, or identify an individual pitch (e.g., middle C) without reference to any other pitch. Whereas relative pitch is the norm among trained and untrained listeners, absolute pitch is rare, occurring in about one in 10,000 people (Levitin & Rogers, 2005; Takeuchi & Hulse, 1993).

Absolute pitch can be a valuable skill for musicians, but it can also interfere with the ability to perceive pitch relations (Miyazaki, 1993, 2004). Because melodies are defined by pitch and duration *relations* and not by absolute pitches, relative pitch is arguably a more musical mode of pitch processing.

Nonetheless, some forms of long-term memory for pitch may be quite common. Halpern (1989) asked participants to sing familiar tunes on various occasions (e.g., "Happy Birthday"), and Bergeson and Trehub (2002) asked mothers to sing the same song to their infant on two occasions. In both studies, singers varied minimally in pitch level across the performances. In another study (Levitin, 1994), undergraduates were asked to sing the first few words of their favorite rock song. Their renditions were very close to the pitch level of the original recording (see also Terhardt & Seewann, 1983). Listeners also exhibit accurate memory for other surface (nonrelational) features of recordings, such as tempo (Levitin & Cook, 1996) and overall sound quality (Schellenberg, Iverson, & McKinnon, 1999). In other words, memory for absolute aspects of musical recordings are more prevalent than previously assumed.

Schellenberg and Trehub (2003) devised a purely perceptual task in which they presented listeners with two versions of a five-second excerpt from an instrumental TV theme song. One version was at the original pitch and the other was shifted upward or downward by one or two semitones. Listeners were better than chance at identifying the original, even in the one-semitone condition. A control experiment confirmed that successful performance was not due to electronic artifacts of the pitch-shifting process.

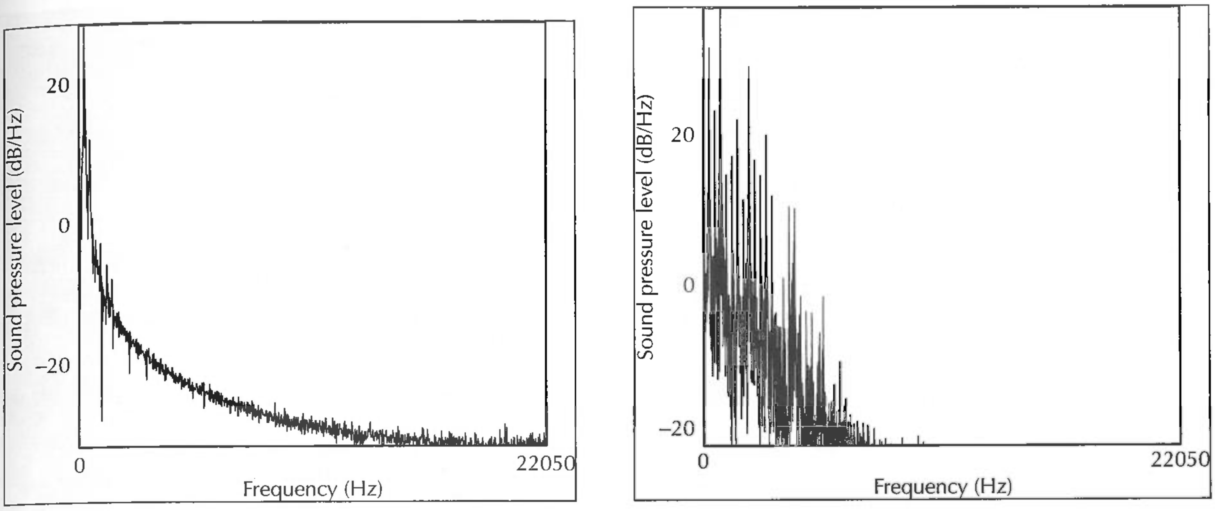
The discovery of good memory for pitch level in adults is pertinent to claims of a developmental shift in pitch processing. The traditional view is that everyone is born with absolute pitch, but this ability disappears unless one receives musical training early in childhood (i.e., by age six or seven). Those without such early training are thought to shift to relative pitch processing, presumably as a consequence of hearing the same pitch relations (i.e., the same tunes, such as "Happy Birthday" or "Twinkle, Twinkle, Little Star") at multiple pitch levels (e.g., sung by a man or a woman). Findings indicating that young infants attend more to absolute than to relative pitch information are consistent with this perspective (Saffran, 2003; Saffran & Griepentrog, 2001).

Nonetheless, in some instances, infants remember pitch relations rather than absolute pitch level (Plantinga & Trainor, 2005). Indeed, there is an abundance of evidence that infants are sensitive to relational pitch properties (contour, interval structure) of melodies (Trehub, Schellenberg, & Hill, 1997). When considered as a whole, the available findings indicate that both absolute and relative pitch processing are evident across the life span. Although relative pitch is easily encoded and accessed from long-term memory, sensitivity to absolute pitch also plays a role in musical experience. In particular, memory for absolute pitch is likely to be evident for brief periods of time, and for stimuli that are heard repeatedly at an identical pitch.

Timbre

Timbre is often described as the attribute distinguishing sounds that are equivalent in pitch, duration, and loudness. Timbre is influenced by the pattern of partials that are present in complex waveforms, and how those partials change over time (i.e., transient or dynamic attributes). A sawtooth waveform, which contains all harmonics, has a timbre that is distinct from a square waveform, which contains only odd-numbered harmonics. The sound of a plucked instrument such as a harp has a relatively rapid amplitude onset, whereas the sound of a bowed instrument such as a violin has a more gradual onset. In naturally occurring sounds, inharmonic partials (i.e., partials with a frequency equal to a noninteger multiple of the fundamental) also affect timbre. If inharmonic partials are removed from a piano note, for example, the note sounds artificial and unfamiliar.

For many instruments, the fundamental frequency has the greatest intensity of all harmonics in the frequency spectrum. Intensity typically decreases for higher-frequency components. An inverse relationship between harmonic number and intensity (called *spectral roll-off*) does not hold for all instruments, however. For some instruments, certain harmonics may be disproportionately intense, giving the instrument its unique timbral character. A "bright" sounding tone, such as that produced by a clarinet, typically contains high-frequency harmonics sounded at relatively high amplitudes.



The frequency spectra for a piano (left) and violin (right); for each, the tone is middle C (261.4 Hz), and the spectrum sample is taken 0.7 seconds after tone onset. The x-axis represents frequency components (starting with the fundamental at the far left), and the y-axis represents intensity. Although intensity peaks at the fundamental for both instruments, the instruments differ in the amount of spectral energy at higher frequencies.

When other aspects of a tone are held constant, differences in the frequency spectrum are associated with differences in timbre. We know, however, that the frequency spectrum is not entirely responsible for timbre, because tones with very different spectra can be perceived as having the same timbre. For example, examination of the frequency spectrum for a note played very softly on a trumpet reveals that most of the partials associated with trumpet sounds are absent. Nonetheless, the note is still perceived as emanating from a trumpet. More generally, the frequency spectra for soft and loud sounds produced by the same musical instrument may be quite dissimilar, despite their perceptual invariance.

Figure 3.3 illustrates three important attributes of a musical instrument that contribute to its perceived timbre. The figure shows the partials that are present, their relative intensity, and how their intensity changes over time. For this particular instrument—a clarinet—the constituent frequency components change in intensity over time at a similar rate. For other instruments, such as a piano, partials change intensity at very different rates. In many instances, onset or "attack" cues (i.e., transient cues during the initial portion of the spectrum) are important for perceiving timbre.

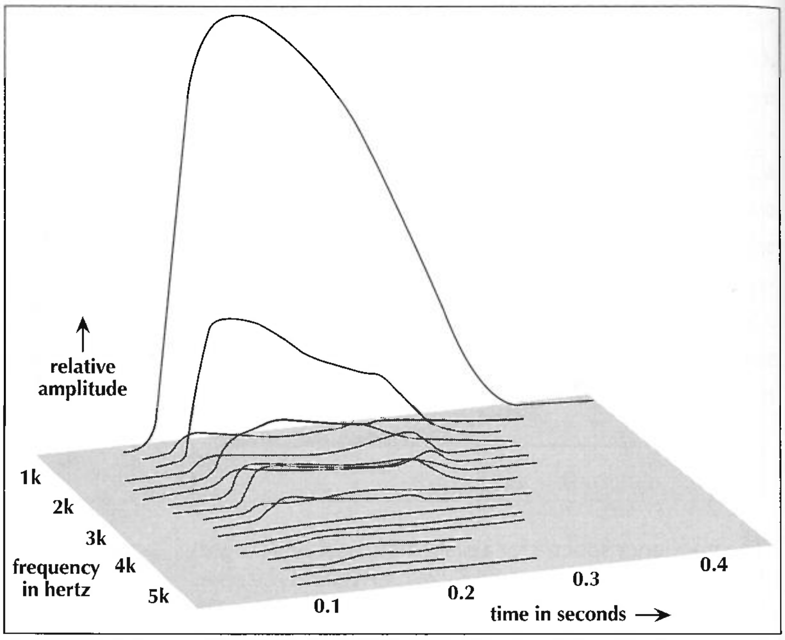
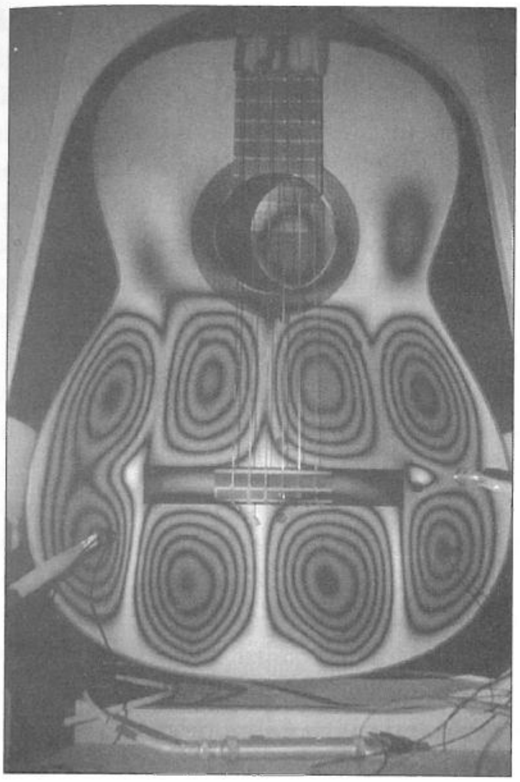


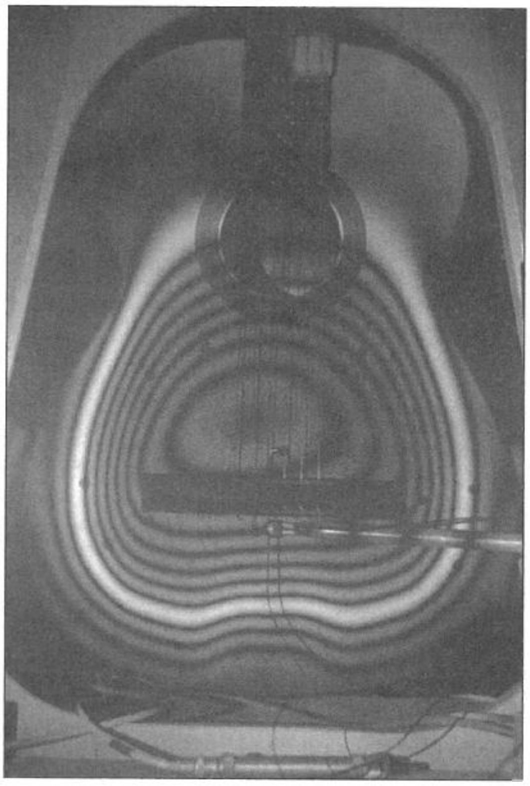
Figure 3.3

Frequency components of a clarinet.

Formants are also thought to influence the perception of timbre. A formant is a range of frequencies with high amplitude relative to other frequencies, which is a consequence of the resonant properties of the sound source (e.g., the body of a piano or an acoustic guitar). Formants correspond to local peaks in a frequency spectrum. Fixed-frequency formants are particularly interesting because they remain relatively constant across the *tessitura* (pitch range) of an instrument. Whereas individual partials of a sound supply information about timbre by their relation to each other and to the fundamental frequency, fixed-frequency formants are resonant frequencies that do not change in proportion with changes in overall pitch. The harmonics that fall within the formant region resonate more loudly than other frequencies, giving the sound (e.g., a musical instrument) its particular timbral character.

Hajda, Kendall, Carterette, and Harshberger (1997) described various methods used to identify the acoustic properties that influence the perception of timbre. These include identification tasks (name that timbre), grouping and classification tasks (match similar





Through the use of an advanced imaging process, the fixed frequency formants of the body of the guitar can be seen as a series of concentric circles; the formant patterns represented in each picture are for two different frequencies.

Source: Photos courtesy of Bernard Richardson, Cardiff University.

sounding timbres), semantic rating scales (rate how bright or dull a timbre sounds), discrimination tasks (are two timbres the same or different?), and various types of similarity ratings (rate how similar two timbres sound). In a classic experiment, Grey (1977) asked musically trained listeners to provide similarity ratings for pairs of notes that came from a set of nine instrumental timbres. Notes were presented with the same fundamental frequency, intensity, and duration so that they varied only in timbre. His analysis revealed that three distinct dimensions could account for a good portion of the variance in similarity ratings.

Various follow-up studies have yielded differing results regarding the underlying dimensions that influence the perception of timbre. In almost all cases, brightness appears to be a significant perceptual dimension. Instruments with a wide range of spectral energy

Sound Example 3.7

Bob Dylan's controversial 1965 concert at the Royal Albert Hall featured electric guitars and by some accounts was met with boos and catcalls by many of his diehard folk fans. These fans had been familiar with his use of a traditional acoustic guitar. **Sound example 3.7** is an excerpt from the concert version of "Like a Rolling Stone." The fact that what ultimately amounts to a change of timbre was met with such hostility suggests a prominent role for timbre in both our musical preferences and our perception of musical styles.

Sound example 3.7 is available via the *Music, Thought, and Feeling* online playlist on www.oup.com/us/thompson.

Sound Example 3.8

is an excerpt of Miles Davis playing a muted trumpet on "Round Midnight." A muted trumpet is an example of high-pass filtering. The mute absorbs many of the lower frequency components, resulting in a tinnier sound.

Telephones and radios are examples of band-pass filtering. In band-pass filtering, lower and higher frequencies are filtered out and a band of frequencies is allowed to pass. Telephones allow only frequencies between 300 and 3,400 Hz. This is a reasonable frequency range for the transmission of speech sounds, because timbral fidelity is not as important in this context as it is in a musical one.

Sound example 3.8 is available via the *Music, Thought, and Feeling* online playlist on www.oup.com/us/thompson.

(e.g., oboe) are perceived as having a brighter timbral character than instruments with a more restricted range of spectral energy (e.g., French horn). The influence of onsets (attacks) and other temporal properties—important in some studies but not in others—depends on the specific timbres used in the stimulus set and the duration of the stimuli (see Hajda et al., 1997).

Another way to identify acoustic attributes that are essential to instrument identification is to examine how readily tone sequences involving different timbres form separate auditory streams. Iverson (1995) presented listeners with sequences consisting of two orchestral tones and asked them to rate the extent to which one or two sequences were being played. Auditory stream segregation was influenced by differences in both static spectra (the pattern of partials present) and dynamic attributes. Moreover, the results converged with similarity judgments obtained by Iverson and Krumhansl (1993) on the same tones. That is, similarity judgments predicted the extent to which timbres segregated into separate auditory streams, with similar timbres less likely than dissimilar timbres to split into separate streams.

How rapidly do listeners register timbre? Robinson and Patterson (1995) tested the limits of listeners' ability to identify the timbre and pitch of tones that varied in duration. Whereas identification of pitch proved to be a function of duration, identification of timbre was independent of duration. Listeners required several complete cycles of a periodic signal to identify pitch accurately, but only one or two cycles (i.e., milliseconds) to identify timbre. These results make intuitive sense when we consider that pitch is determined by frequency (number of cycles per second). A listener needs to hear several cycles to determine how rapidly a cycle repeats. By contrast, even one or two cycles of a tone contain information about the components present in the frequency spectrum, which allows rapid identification of timbre.

In a related investigation, researchers tested the limits of listeners' ability to identify songs (Schellenberg et al., 1999). In an experimental version of *Name That Tune* (the song-identification game played on TV and radio), listeners matched extremely brief excerpts (100 or 200 milliseconds, or msec) from recordings of popular songs with the song titles and artists. Listeners' performance was better than chance with excerpts of 1/5 of a second. Performance deteriorated with shorter excerpts (1/10 of a second) but was still better than chance. Other manipulations involved low-pass and high-pass filtering, and playing the excerpts backward. Performance fell to chance levels for the low-pass filtered and backward excerpts but was unaffected by high-pass filtering.

In short, successful identification of the recordings required the presence of dynamic, high-frequency spectral information. Note that the excerpts were too brief to convey any relational information (e.g., pitch and duration), and the absolute pitches of the excerpts would be identical whether they were played backward or forward. Thus, successful performance presumably stemmed from accurate and detailed memory of the recordings' timbres, which consisted of many complex tones from many instruments.

By definition, timbre and pitch are different. Nonetheless, some studies suggest that timbre and pitch are not perceived independently. For example, it is easier to judge whether two tones (Crowder, 1989), two chords (Beal, 1985), or two melodies (Radvansky, Fleming, & Simmons, 1995) are the same when standard and comparison stimuli are played on identical rather than different musical instruments. This finding suggests that despite our experience of pitch and timbre as distinct musical qualities, they interact with each other in music processing.

Krumhansl and Iverson (1992) used another approach to study interactions between pitch and timbre. Participants were asked to classify stimulus items into two groups as quickly as possible on the basis of a specified target attribute (e.g., pitch). Not only did the target attribute vary; an irrelevant attribute (e.g., timbre) also varied from trial to trial. Performance on the categorization task was relatively good when variation in the irrelevant attribute was correlated with variation in the target attribute, but relatively poor when variation in the irrelevant attribute was uncorrelated with variation in the target attribute. That is, participants were unable to attend to the pitch of a tone without being affected by its timbre, or vice versa. This finding yields converging evidence that pitch and timbre are not perceptually separable from each other.

Although pitch and timbre may be integrated successfully at some points in the information processing system, listeners can also become confused about how pitches and timbres are combined. In one experiment, participants listened for particular combinations of pitch and timbre in arrays of tones that were presented simultaneously but emanated from a number of locations in the auditory field

Sound Example 3.9
is Debussy's "La Cathédrale Engloutie" (The Sunken Cathedral). The section from 2:27 to 3:13 is an example of the timbral effect of tonal fusion. It is difficult to pick out a prominent melody; the constituent tones of each chord are perceptually fused.

Sound example 3.9 is available via the *Music, Thought, and Feeling* online playlist on www.oup.com/us/thompson.

Sound Example 3.10.1–3.10.2

These are two examples of "timbre for timbre's sake." In both of these compositions, one acoustic and one electronic, timbre is used as a compositional end in itself. **Sound example 3.10.1** is the second movement ("Hawthorne") from Charles Ives's *Concord Sonata* (1909–1915) for piano. At 1:22, as indicated in the musical score, the pianist depresses a two-octave span of the keys using a 14 3/4" board at various spots across the keyboard.

Sound example 3.10.2 is "Where the Wind Meets the Sea" (1999) by Timothy Opie, in which timbral elements are electronically synthesized. This piece is composed of tiny "grains" of sound that are less than 40 msec in length—a compositional technique known as granular synthesis.

Sound examples 3.10.1 and 3.10.2 is available via the *Music, Thought, and Feeling* online playlist on www.oup.com/us/thompson.

(Hall, Pastore, Acker, & Huang, 2000). An examination of errors indicated that participants often perceived an illusory conjunction of pitch and timbre. That is, participants often heard the timbre of one tone combined with the pitch of another tone. Estimates of illusory conjunction rate ranged from 23% to 40%. The findings are evidence that after an initial stage in which individual features of musical tones are registered (e.g., pitch, timbre), there is a stage in which these separately registered features are integrated (feature integration). Illusory conjunctions arise when errors occur at the feature integration stage of processing.

What is the relation between timbre and harmony? The principle of *harmonicity*, outlined by Bregman (1990), suggests that partials falling along the harmonic series of a single note are fused together, and so listeners perceive the distinct timbral qualities of each note holistically. When two instruments play the same note, or two notes with many harmonics in common, we may have difficulty telling the two notes apart. Such fusion effects actually contribute to our perception of harmony. For example, notes separated by a fifth have several partials in common. When they are played at the same time, perceptual mechanisms tend to fuse or "confuse" some of the components of the two notes, giving rise to emergent timbral effects. Compositions by the French impressionist composer Claude Debussy exemplify the varied timbral effects of combining tones.

Although pitch and rhythm are the most basic elements of a musical composition some contemporary composers and theorists have considered the possibility that timbre might be used compositionally in a way that is analogous to the use of pitch (e.g., Lerdahl, 1987; Slawson, 1985). One challenge with this compositional approach is that timbre has a powerful influence on auditory stream segregation. Thus, any large shift in timbre is likely to signal a separate auditory stream, which would conflict with the goal of creating a coherent sequence analogous to a melody.

Nonetheless, changes in timbre that correspond to reductions in sensory dissonance (i.e., fewer harmonics falling within a critical band)—and increases in fusion—may be perceived as a change from tension to release, which commonly occurs at points of musical closure (Pressnitzer, McAdams, Winsberg, & Fineberg, 2000). In other words, changes from relatively dissonant to relatively consonant timbres have some of the resolving properties of a perfect cadence. Moreover, it is clear that timbre has a profound influence on our appreciation of music. For example, imagine the music of Radiohead played on accordion, or a rock song by U2 played on an oboe. Such changes in instrumentation would undoubtedly affect our experience of the music.

Reconstructing Music

Sensory processing is initiated when the eardrum responds to the vibrations of air molecules. This sensory input is analyzed into a number of attributes or features, among them pitch, timbre, and intensity. The same analysis occurs for all types of sound, whether it arises from wind, rain, music, or a barking dog. Sensory analysis, in turn, is just one link in a complex chain of neural events that ultimately give rise to a full auditory experience. Among those processes are ones that help to create a unified perception, so that the various attributes of sound are not experienced independently of one another. At some levels of processing, the pitch and temporal structure of music are processed in different parts of the brain, but listening to music does not generate independent experiences of pitch and rhythm. Rather, the dimensions of music seem to be merged into an integrated and unified experience. To understand these unified experiences requires examination of multiple levels of music structure and brain processing: from the sensory building blocks, to our perception of melody and rhythm, to our deepest emotional responses, to the skilled performances of experts, and to all of the creative activities associated with music.

Additional Readings

Mathews, M. (1999). The ear and how it works. In P. R. Cook (Ed.), *Music cognition and computerized sound: An introduction to psychoacoustics* (pp. 1–10). Cambridge, MA: MIT Press.

Oxenham, A, J. (2013). The perception of musical tones. In D. Deutsch (Ed.), (2013), *The psychology of music* (3rd ed., pp. 1–34). San Diego: Academic Press.