13 Pitch

The word sonority *encapsulates aesthetic and structural aspects of tone simultaneities in MmT. Related to* texture*, sonority is a subjective property of musical tones, chords, passages, and pieces. The perception of tone simultaneities depends on the perception of individual tones. The pitch of a pure tone or partial is* spectral*, whereas that of a complex tone (voiced speech sound, musical instrument, environmental sound) is* virtual*. A virtual pitch does not correspond to a specific partial but is derived from multiple spectral pitch contours by spectral/temporal pattern recognition. A virtual pitch is heard when audible partials describe an incomplete harmonic series as if they came from the same periodic source. The main virtual pitch of a harmonic complex tone in the central musical range is determined by higher harmonics; it is not shifted by shifting the fundamental. Since virtual pitch may correspond to both temporal and spectral aspects of a signal, it is difficult to predict from physical measurements of neural activity. A pattern-recognition model of pitch perception can explain the nature and origin of musical chord roots and tonics.*

Measuring Pitch

Frequency and pitch are very different things. The frequency of a periodic sound is the number of vibrations per second. The pitch of a sound is its perceived height on a one-dimensional scale from low to high. In everyday life, the number of times per second something happens has nothing at all to do with the height of something above the ground. A periodic sound does not in any physical sense have height; nor can one perceive individual vibrations within a regular steady tone in the range of human hearing (20 Hz–20 kHz).

It is clear how frequency is measured in physics. Count how often a periodic pattern repeats per unit of time. If the frequency is constantly changing, limit the counting to a temporal window and count within several overlapping windows.

Pitch is not so easy to measure. Like any other quantitative psychological variable, it needs to be operationalized in a standard experimental procedure. Different experimental measures of pitch can only be compared if the method is standardized. Misunderstandings in the pitch-perception literature may arise from differences in experimental methods; sometimes, those differences involve unstated philosophical assumptions.

Pure tones are appropriate standard references for pitch experiments because they are physically simple. There are few parameters to adjust and few psychoacoustic dependencies. The pitch of a pure tone is relatively clear and unambiguous (pitch salience is relatively high). Whereas the timbre of a pure tone changes with frequency (from a low rumble to a high squeak), the pitch–timbre relationship is always the same for a pure tone of given amplitude and a given listener.

In a typical standard procedure, a listener alternately hears a short test sound and a short pure reference tone, with short time gaps between them, and adjusts the frequency of the pure tone until the pitch of the test sound matches the pitch of the pure tone (Terhardt 1988, 312). Here, “short” means a few hundred milliseconds (typical of speech and music), and the SPL of the pure tone is understood to remain constant. In this way, the pitch or pitches measured in a sound can be measured in equivalent hertz.

In such experiments, psychoacousticians traditionally ignore individual differences. Instead, they average results over several listeners on the assumption that individual differences are relatively small. Perhaps they should not, given differences between spectral and fundamental listeners (Seither-Preisler et al. 2007).

Using this method, one can show that the pitch of a pure tone depends not only on its frequency, but also on its intensity and on the presence of other simultaneous sounds (pitch shift). One can also show that the pitch of a harmonic complex tone is ambiguous: depending on how tones are presented in a random sequence, listeners may match the pure reference tone to different pitches within the same complex. Those pitches do not necessarily correspond to partials.

Different Kinds of Pitch

Some music psychologists regard pitch as comprising two psychological dimensions—pitch height and pitch class—but the idea is problematic (see chapter 14). Instead, we can distinguish two kinds of pitch: spectral and virtual (Terhardt 1972). In a complex tone (i.e., a tone comprising different partials of different frequencies), spectral pitch is the pitch of a single partial, and virtual pitch is the pitch of the whole tone.

When we hear a pitch, it is not immediately apparent whether it is spectral or virtual. Both spectral and virtual pitches are associated with diverse timbres (using the word “timbre” in the sense of a subjective experience—what a tone sounds like). We may for example hear a church bell with two pitches: a strike tone with a virtual pitch, and a hum tone with a spectral pitch (Terhardt and Seewann 1984). Without theoretical knowledge about strike and hum tones, it is unclear which is which. The terms “spectral” and “virtual” refer to the origin of the pitches—not what they sound like. They refer to what each pitch depends on and how it might be predicted. Spectral pitches depend directly on spectral frequency, whereas virtual pitches involve pattern recognition.

Rather than distinguishing two kinds of pitch, it is sometimes more intuitive to refer to two kinds of tone sensation. A pure tone sensation is the experience of a listener who is consciously perceiving a pure tone; a complex tone sensation is the experience of a listener who is consciously perceiving a complex tone. A spectral pitch is the pitch of a pure tone sensation, whereas a virtual pitch is the pitch of a complex tone sensation (Parncutt 1989).

In a typical harmonic complex tone, there is both a virtual and a spectral pitch at the fundamental. We are typically aware only of the virtual pitch because it is more salient. The tone’s (virtual) pitch and timbre usually depend on the upper partials more than they depend on the fundamental. If the harmonic complex tone has no fundamental, the main virtual pitch typically corresponds to the missing fundamental.

An everyday example is a voice heard on the telephone. Telephones usually do not transmit acoustic frequencies below about 300 Hz. When the phone rings and we answer, not expecting anyone in particular, we quickly assess whether the speaker has a higher voice (probably a woman or child) or a lower voice (probably a man). To make that assessment, we perceive the virtual pitch at the missing fundamental. In everyday nonmusical perception, most pitches are virtual and perceived in the same way (Terhardt 1974a). The telephone example illustrates that virtual pitches are a normal, effortlessly perceived part of everyday life.

Spectral Pitch

From the viewpoint of acoustics and the auditory scene, the simplest sound is a pure tone, in which the relationship between air pressure and time is sinusoidal. Spectral pitch is the pitch of a pure tone, or the pitch of an audible partial within a complex tone when the tone is perceived analytically.

Spectral pitches are subjective experiences of the listener, whereas spectral frequencies are physical measures. Spectral pitches and spectral frequencies differ in two ways: small deviations called pitch shifts, and variations in audibility. Masking can cause a partial to become aurally irrelevant: the spectral frequency exists but the spectral pitch does not.

Terhardt et al. (1982b) assumed that a spectral pitch is perceptually salient, and hence more likely to be noticed separately ("heard out") or to contribute to virtual pitch, for three reasons:

- It has relatively high SPL relative to masked threshold. That's how overtone singers make one overtone in their voice audible: by changing their vocal tract resonances to increase the sound level of that specific overtone. If the SPL is already well above masked threshold, increasing it further will increase loudness but not pitch salience.
- It lies near the center of the frequency region within which the ear is most sensitive to spectral pitch. The spectral dominance region lies between about 300 and 2,000 Hz, with a peak near 700 Hz; it corresponds roughly to the range of the first two speech formants (cf. B. C. J. Moore et al. 1985; Ritsma 1967; Terhardt et al. 1982b). S. S. Stevens et al. (1947) estimated that the frequency region for speech transmission is confined to 250–7,000 Hz.
- It is not masked, having a relatively high SPL or being relatively distant in frequency from other partial tones or noises.

Virtual Pitch

A virtual pitch is the pitch of a complex tone when perceived holistically. In Gestalt psychology, most perception in everyday life is holistic. In an approach called holism, perceptual experiences are regarded as intrinsically holistic (Wagemans et al. 2012).

Virtual-pitch perception involves pattern recognition in the auditory scene. The physiological foundations are complex, involving a mixture of innate and learned processes, and temporal and spectral patterns (Assmann and Summerfield 1990; Vliegen et al. 1999). But the same applies to the pitch of a pure tone: it also depends on both spectral and temporal information (B. C. J. Moore 1989; Oxenham 2013). The neural foundations of virtual pitch are located in the auditory cortex (Bendor and Wang 2005; Lütkenhöner et al. 2010; Ross et al. 1996)—specifically, Heschl's gyri (Krumbholz et al. 2003; Zatorre 1988); Catherine Liégois-Chauvel and colleagues (2001) and Andrew Oxenham and colleagues (2004) investigated cortical tonotopic representations. For music-theoretical purposes, we can ignore the physiology of virtual pitch, which is complex and poorly understood, and focus instead on the listener's subjective auditory experience.

Helmholtz (1856) clarified the physical difference between pure tones (*einfache Töne*) and complex tones (*Klänge*) and between combination tones and beats, reconciling differences between Seebeck and Ohm (Steege 2012). A pure or complex tone is something that exists in sound before it hits the ear. Combination tones and beats are created by the ear, and they are two quite different phenomena. Combination tones are distortions that produce new frequencies that can be heard as new pitches. The vibrations exist as additional pure tones on the basilar membrane, even if they do not exist in the sound outside the ear. Beats are amplitude modulations; beat frequencies do not evoke pitch. A pure tone whose amplitude is varying has a pitch corresponding to the carrier frequency; the beating is perceived as variation in loudness. If the modulation is fast, the individual beats can no longer be heard, and the sound becomes rough.

The relationship between spectral and virtual pitch normally has nothing to do with combination tones (which are inaudible at usual sound levels for both speech and music; Plomp 1965) or beats (which do not evoke new pitches). The relationship is difficult to investigate neuroscientifically because both spectral and virtual pitch belong to World 2, and the relationship between Worlds 1 and 2 is the famously intractable mind-body problem.

Helmholtz explained why we are usually unaware of spectral pitches in everyday sounds: "Whether the voice of a dog contains the higher octave or the twelfth of the fundamental tone is without practical interest and not an object of our attention. Thus the overtones are assimilated into those peculiarities of a tone, not to be further specified, which we call its timbre" (Steege 2012, 75; translated from Helmholtz 1858). Put another way, our sensory organs are constantly sending information to our brain, but we only become aware of a small fraction of that information. Those percepts of which we are consciously aware tend to correspond to environmental affordances, and their pitches are virtual.

To understand speech, the brain must track the fundamental frequency, allowing the listener to focus on the speech signal and ignore the background noise. The fundamental frequency contour communicates prosodic meanings that augment the lexical meaning of the words being spoken. To follow it, the brain needs to recognize harmonic patterns in the running spectrum of the auditory scene and separate them from the background noise.

We might therefore predict that the most salient virtual pitches lie in the fundamental frequency range of speech and singing (about 65–1,000 Hz). Calculations using the algorithm of Terhardt et al. (1982b) are consistent with that prediction (Huron 2001). The range is close to that of Bach's clavichord (C2–C6), and virtual-pitch salience peaks near about 300 Hz (D4). The most salient spectral pitches lie in a higher region that is centered on about 700 Hz and roughly symmetrical in log frequency.

The (virtual) pitch of a harmonic complex tone is generally ambiguous. In holistic perception, it may be an octave higher than the fundamental, the even-numbered harmonics being perceived as a tone within a tone. To demonstrate the effect using a digital sound editor, slice a high note out of an operatic tenor solo. Let's say the note is A4, which is theoretically 440 Hz; in fact, it may be a little higher because solo intonation tends to be slightly sharp relative to the accompaniment (Kantorski 1986). Now, compare the pitch of the tenor's voice in the recording with two successively presented pure tones: one at the same fundamental frequency (say, 445 Hz) and one an octave higher (say, 990 Hz). If both seem to match the original quite well, which is *the* pitch? That is one kind of ambiguity, and there is another rather simple one: if a complex tone is heard analytically, individual partials can be heard out (the spectral pitches). Again, it is unclear which is *the* pitch.

It is possible to hear a pure tone sensation and a complex tone sensation at the same time with the same pitch. In other words, a spectral pitch can coincide with a virtual pitch, and we can be aware of the coincidence. It can happen when our attention is attracted to the fundamental of a regular harmonic complex tone—the pure tone at the fundamental frequency. Imagine a sound demonstration in which a complex tone alternates with a pure tone and the frequency of the pure tone is gradually changed (Houtsma et al. 1987; Terhardt 1998). As the pure tone approaches each partial, we start to hear that partial as a separate tone within the complex tone—we "hear out" the harmonic. The same thing happens when the pure tone approaches the fundamental: we start to hear the pure tone at the fundamental as a different tone with the same (spectral) pitch as the (virtual) pitch of the complex tone as a whole but a different timbre (that of a pure tone sensation rather than a complex tone sensation). The demonstration confirms that there are two pitches at the fundamental: a spectral pitch corresponding to the partial at the fundamental, and a virtual pitch corresponding to the tone as a whole associated with a different timbre.

Another way to demonstrate this effect is to present two harmonic complex tones in alternation that are identical except that the fundamental is missing from one of them. The comparison attracts attention to the fundamental when it appears in the complete tone. We hear the fundamental as a separate pure tone—with the same pitch as the complex tone but different timbre.

Virtual pitches can be evoked by patterns of unresolved higher harmonics—partials that are inaudible in the sense that listeners cannot attend to them. The effect can be demonstrated by presenting a well-known melody from tones comprising only high harmonics that cannot be resolved and asking a listener to name it (Moore and Rosen 1979). A monist—someone who believes that only the physical world exists—might argue that the spectral pitches in question do not exist, but the virtual pitch does. A dualist, who also believes in the existence of subjectivity, might come to a different conclusion. If a spectral pitch affects subjective experience in any way, it is aurally relevant in that sense. Aurally relevant partials form a spectral-pitch pattern that, like the spectral pitches it contains, belong to World 2—even if the listener is incapable of analyzing the pattern into its parts (i.e., even if the spectral pitches are unresolved).

The results of Roelof Ritsma (1967) and others are consistent with the following principle: A tone's most salient virtual pitch is a subharmonic of its most salient spectral pitch(es). Consistent with that idea, Houtgast (1976) found that even the pitch of a pure tone sometimes corresponds to a subharmonic, especially if the pure tone is presented within noise that could have masked a tone at the subharmonic. Dik Hermes (1988) developed a pitch model based on subharmonic summation, and the success of Terhardt's pitch model is further evidence for this principle. Similarly, the root of a chord tends to correspond to a subharmonic of one or more chord tones.

The predicted spectral and virtual pitches of an isolated pure tone on C4 and an isolated harmonic complex tone on C4 according to these principles

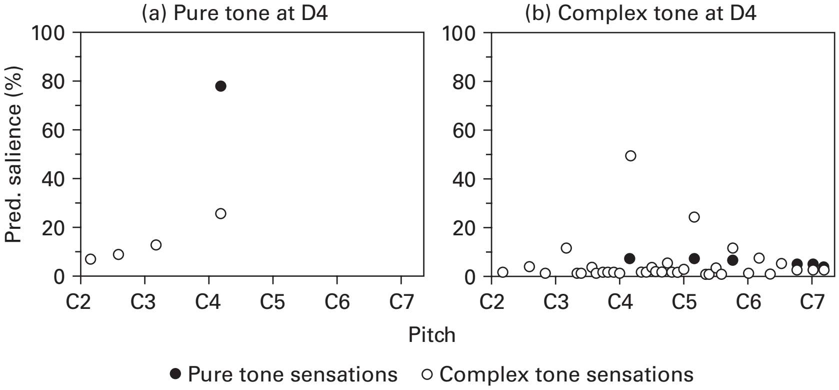

Figure 13.1

Pitches evoked by (a) a pure tone and (b) a harmonic complex tone according to Parncutt's (1989) adaptation of the pitch algorithm of Terhardt et al. (1982b). Filled circles are spectral pitches (pure tone sensations), and open circles are virtual pitches (complex tone sensations). The spectrum of the harmonic complex tone on the right is musically typical (the amplitude spectrum is specified in Parncutt 1989). Pitch-salience values have been scaled to predict the probability of noticing.

are shown in figure 13.1. Part (a) of the figure features weak subharmonic virtual pitches of a pure tone that have been observed empirically in the presence of background noise (Houtgast 1976) and are not predicted by Terhardt et al. (1982b). The pitch ambiguity of the complex tone (in particular, at octave and 5th intervals) is consistent with empirical studies on the similarity of successive pure and complex tones, including Parncutt (1989), Gerhard Stoll and Parncutt (1987), and Terhardt et al. (1986). Rising octave intervals between successive harmonic complex tones are more similar than falling, possibly due to an implication–realization relationship (similar to that described by Narmour 1990): the second-most-salient pitch in the first sound is realized as the most salient pitch in the second sound. For the same reason, *falling* octave intervals between successive *pure* tones are more similar than rising, as shown in figure 13.2.

Spectral-Pitch Shift

The perceived pitch of a pure tone (how high it sounds) can be shifted slightly (so that it becomes slightly sharp or flat) without changing its frequency, in

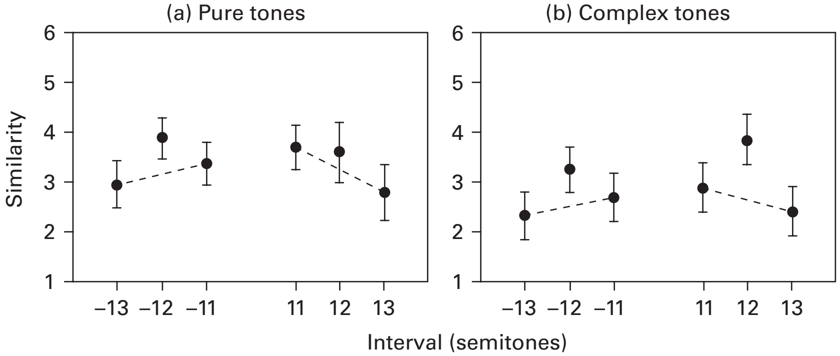

Figure 13.2
Mean similarity judgments of (a) pure and (b) complex (sawtooth) tones spanning major-7th, octave, and minor-9th intervals (eleven, twelve, and thirteen semitones). Rising intervals are assigned positive values on the horizontal axis; falling intervals are assigned negative values. Dots are mean similarity ratings over thirty listeners. Error bars are 95 percent confidence intervals. Adapted from Parncutt (1989, figure 5.5b).

two ways: by changing its amplitude (the tone seems to go slightly sharp or flat if suddenly made much louder or quieter), or by partially masking the pure tone with other, simultaneous sounds (the tone's pitch moves away slightly from the pitch of the masker). Pitch shifts of this kind have also been called pitch bias (e.g., Chuang and Wang 1978).

Many empirical studies have independently confirmed and measured spectral-pitch shifts (Cheveigné 1999; A. Cohen 1961; Hartmann and Doty 1996; Morgan and Galambos 1943; W. B. Snow 1936; S. S. Stevens 1935; Stoll 1985; Verschuure and Van Meeteren 1975; Walliser 1969). Whereas individual differences have been observed in the size of the shifts, the direction of the shift is usually the same for different listeners.

Despite repeated empirical confirmation of this phenomenon, some researchers remained skeptical. Robert Peters and colleagues (1983) had three listeners match the pitches of partials in harmonic complex tones comprising harmonic numbers 1–7 or 4–7 (the latter having a missing fundamental). Not expecting to find shifts, they found that "the pitches of components in context were again found to be very close to those of components in isolation" (abstract), implying the existence of small deviations.

Spectral-pitch shifts originate in the physiology of the cochlea—the complex and imperfect way in which it generates a running spectral analysis over a very wide range of sound levels (Evans 1981). The cochlea has no easy task: a difference of 120 dB between the quietest and loudest perceptible sound corresponds to a factor of one million in air pressure and one trillion in power. From that viewpoint, it is unsurprising that there are pitch shifts. In fact, it is surprising that pitch shifts are as small as they are. In an ecological-evolutionary approach, the reason they are so small involves the importance of frequency patterns for the recognition of environmental sound sources (see chapter 6).

We can confirm the existence of pitch shifts informally. Listen to a pop song that starts with a solo bass line. In the terminology of cognitive music psychologists such as Krumhansl (1990), the solo bass line instantiates the song's tonality in the mind of the listener. During the introduction, suddenly and drastically change the loudness. Compare the pitch at very low sound level with the pitch at very high sound level. Alternatively, listen to a solo bass line at a low sound level, although it was originally recorded at a high level. The pitch shift becomes clear when a higher instrument or voice enters. The following original recordings may serve as examples: *My Girl* and *Papa Was a Rolling Stone* by The Temptations, *Pumped Up Kicks* by Foster the People, and *Could You Be Loved* by Bob Marley.

Many people have experienced pitch shifts without realizing it. Imagine trying to sleep with a party going on in another apartment or down the road. Mostly, you hear only the bass line: boom, boom, boom. The pitch of those booms allows you to imagine the tonality. Suddenly, someone opens a door, and you start to hear the melody. It sounds like the singer is singing in a different key. That is because you are hearing the bass line more quietly than the people in the house are hearing it. For the people in the house, the music is loud and the bass sounds in tune. For you, the music is quiet, and the bass sounds out of tune.

Virtual-Pitch Shift

Virtual pitches may also be shifted, but not as much as spectral pitches (Stoll 1985)—perhaps because a virtual pitch typically depends on several aurally relevant partials (spectral pitches), only some of which are shifted by changing SPL or introducing maskers. Virtual pitches may be shifted for two reasons. First, they depend on shifted spectral pitches; if a virtual pitch is a subharmonic of a spectral pitch, and the spectral pitch is shifted, the virtual pitch should be shifted by the same amount. Second, the subharmonic intervals that are applied in Terhardt's pitch model are slightly stretched, because the spectral pitches in the harmonic spectra from which they are learned are shifted.

The second idea can explain why octaves in music are stretched relative to a 2:1 ratio. The octave intervals within the auditory system's harmonic template (between harmonics 1 and 2, 2 and 4, 3 and 6, and so on) are stretched. As a result, the interval between a salient spectral pitch and a corresponding virtual pitch is generally not exactly harmonic. The pitch of a harmonic complex tone with a missing fundamental is often slightly flat relative to the pitch of a pure tone at the same frequency (Terhardt et al. 1982b).

There are also pitch shifts in complex tones with physically stretched harmonics, such as piano tones. In that case, the shifts are both physical and perceptual. If such a tone is perceived holistically, in the usual way, the perceived pitch may change when attention is attracted to the lowest partial.

Experiments on pitch shifts in complex tones are sensitive to variations in method. Stoll (1984) and Allan Vurma and colleagues (2010) claimed that a sung speech vowel with a brighter timbre (e.g., /i/) has a slightly higher pitch than a harmonic complex tone with a duller timbre (e.g., /u/). In each experimental trial, two vowels were presented in succession, and the listener was asked which was higher in pitch. But there was a methodological problem: the tones being compared varied in both pitch and timbre. If the aim was to measure pitch, the results were confounded by timbre, and asking the participants to "ignore any differences of timbre between the stimuli of a pair" (Stoll 1984, 139) could hardly solve the problem. The solution is to hold the timbre of the comparison tone constant, or only to consider trials in which the comparison tone was constant (Chuang and Wang 1978).

No current model of pitch perception reliably and generally predicts pitch shifts of pure tones as a function of intensity and masking, and corresponding shifts in virtual pitches. Terhardt et al. (1982b) claimed to predict the direction of such pitch shifts but not their magnitude. That is reasonable, given large individual differences in the perception of pitch shifts. Temporal models of pitch perception have difficulties dealing with mistuned partials, such as octaves that are not exactly 2:1. Hartmann and Doty (1996) accounted for perceptual pitch shifts in physically shifted (mistuned) harmonics by a combination of spectral (excitation pattern interaction) and temporal (neural timing) effects, but their solution was not generalized to accept any spectrum as input.

Mistuned Harmonics

The literature on pitch shift is confused by different usages. In the present book, and following Terhardt, I use the term "pitch shift" only for small changes in the (spectral) pitches of pure tones as a function of SPL or masking and corresponding changes in the (virtual) pitches of complex tones. As we will see in this section, other researchers used the term differently.

Inspired by Schouten's (1940) theory of the residue, De Boer (1956) investigated the pitch of complex tones whose partials are equidistant on a linear frequency scale but not harmonic. He investigated the shift in virtual pitch when the frequencies of the partials of a harmonic complex tone are increased linearly and found that "midway between two harmonic situations, the pitch is indeterminate" (536). For example, if a complex tone with three partials at 800, 1,000, and 1,200 Hz is shifted upward so the spectral frequencies become 850, 1,050, and 1,250 Hz, the virtual pitch at the missing fundamental rises from 200 to about 210 Hz. In that case, the first tone is harmonic, but the second is not; the auditory system tolerates the small deviation from harmonicity. If the frequencies are further increased to 900, 1,100, and 1,300 Hz, the pitch becomes ambiguous: the three partials could be perceived either as harmonic numbers 4, 5, and 6 of $f_0 \approx 1{,}100/5 = 220$ Hz, or as harmonic numbers 5, 6, and 7 of $f_0 \approx 1100/6 = 183$ Hz. Schouten et al. (1962) showed that the virtual pitch of such approximately harmonic complex tones could not be due to difference tones. He termed the difference between the perceived pitch and the constant spacing of the partials (here, 200 Hz) the first effect of pitch shift. The effect can be heard on the Auditory Demonstrations CD of Houtsma et al. (1987, Demonstration 21 "Shift of virtual pitch").

These results suggest that a virtual pitch is generally an approximate subharmonic of a salient partial. Here, $850/4 = 213$, $1050/5 = 210$, and $1250/6 = 208$. A virtual pitch corresponding to 210 Hz is approximately subharmonic to all three spectral pitches. In a spectral pattern-recognition approach, a spectral-pitch pattern does not have to be exactly harmonic to be recognized by the auditory system as harmonic. A temporal approach can also allow for deviations, but less directly.

These ideas contributed to the formulation of the pitch algorithm of Terhardt et al. (1982b). Terhardt and Manfred Seewann (1984) used the same idea to predict the pitches of church bells, whose spectra are nonharmonic but comprise approximately harmonic subsets. Consistent with pitch theory, J. H. (Berry) Eggen (1986) found that the strike note of a bell depended on the octave (2nd harmonic), the 12th (3rd harmonic), and the upper octave (4th harmonic).

There is also a second effect of pitch shift, in which the pitch of a harmonic complex tone becomes slightly lower if the tone is made slightly inharmonic by stretching the spectrum. Mike Greenhough and colleagues (1973) found that “the magnitude of the second effect depends on the loudness of the stimulus” and “the magnitude of the effect is a function of the position of stimulation of the basilar membrane” (277).

Effects of this kind may be explicable by both time- and frequency-domain models. In Terhardt’s approach, which can be described as a spectral-pitch-domain model (spectral pitches being subjective, whereas frequencies are objective), virtual-pitch shifts depend on spectral-pitch shifts, which in turn depend on SPL and mutual masking. They also depend on learned pitch distances between partials in harmonic complex tones such as speech vowels.

Yet another kind of pitch shift (actually, a frequency shift) involves physically mistuning a partial within a harmonic complex tone. Empirical studies have shown that a mistuned harmonic contributes to the virtual pitch at the fundamental for mistunings of up to 3 percent or half a semitone (Darwin et al. 1994; Lee and Green 1994; B. C. J. Moore et al. 1985). That is unsurprising in an approach based on the learning and recognition of pitch patterns within speech and environmental sounds; in everyday life, the pitches of partials are only approximately defined and can be shifted by changes in amplitude and masking. Moreover, some quasi-harmonic sounds have physically stretched partials relative to a harmonic series (piano tones being an example).

These uncertainties suggest that the brain uses a kind of harmonic template to recognize harmonic complex tones. Given that the elements of the template are only approximately defined, the expression “harmonic sieve” is appropriate (Roberts and Brunstrom 1998). These ideas are merely images that help us to understand the process; what actually happens in the physical world is largely concealed within biological neural networks.

Experiments on the perception of mistuned harmonics are relevant for music theory. The 7th partial of a harmonic complex tone (7th harmonic) is about 1/3 of a semitone flat relative to a minor-7th interval, when all semitones are equal. If the 7th harmonic is raised 1/3 of a semitone to align with 12-EDO, it still contributes to the virtual pitch at the fundamental. In that regard, the difference between a Mm7 chord in 12-EDO and the so-called natural 7th chord (4:5:6:7) is smaller than music theorists of the past have assumed. The natural 7th may sound out of tune to our 12-EDO-influenced ears, but the 7th still contributes to the clarity of the root. Tuning and intonation are in any case inherently and generally approximate (see chapter 7).

There are additional implications for music theory. A musical chord may fuse (or almost fuse) if it is almost the same as a major triad or Mm7 chord, one voice being shifted up or down by up to a semitone. Applying this principle systematically yields a list of chords in which fusion is relatively strong, but not quite as strong as in major triads and Mm7 chords. Shifting the root of a major triad up a semitone produces a diminished triad; shifting it down a semitone produces a minor triad. Shifting the 3rd up or down produces a suspended-4th chord or minor triad, respectively. Shifting the 5th produces a major triad with diminished 5th (tritone) or an augmented triad. All such chords happen relatively often in tonal music (the last two being less common for other reasons: the dissonance of both the tritone and the major-2nd interval between the major 3rd and diminished 5th above the root, and the root ambiguity of the augmented triad). Shifting the tones of an Mm7 chord in a similar way produces chords such as minor 7th, major 7th, diminished 7th, 7th with suspended 4th, 7th with diminished or augmented 5th, and others. This procedure (which is similar to a procedure proposed by Rameau) produces the main sonorities of MmT, with some exceptions (e.g., the half-diminished 7th, which can be created by shifting *two* tones of a Mm7 chord through a semitone). Of course, the degree of fusion of a musical chord also depends on other factors, such as voicing (chordal inversion, spacing, doubling), musical context (immediately preceding and following sounds), and listener (some are more likely to listen holistically, others analytically).

Temporal or Spectral?

Both spectral and virtual pitches, and the relationship between them, depend on both temporal and spectral aspects of sound, or of neural responses to sound. Temporal aspects are aspects that can be seen in the waveform, such as period, waveform shape, changes in the waveform from one repetition to the next (hence, departures from periodicity), and amplitude modulation (beating). Spectral aspects can be seen in the running spectrum or spectral representation—primarily, the changing frequencies and amplitudes of the partials and noise components. Although several researchers have compared the performance of spectral versus temporal models of pitch or have assumed that one of the two is superior to the other (M. A. Cohen et al. [1995] preferring a spectral model, Patterson et al. [1995] a temporal model), results of such comparisons have been indecisive—in part due to the mathematical equivalence of periodicity and harmonicity.

From an evolutionary or ecological viewpoint, pitch perception is constrained both physiologically and environmentally: "pitch perception is critically shaped by the constraints of natural environments in addition to those of the cochlea" (Saddler et al. 2021, 1). The brain therefore takes advantage of both spectral and temporal cues (B. C. J. Moore 1989; Klapuri 2004): the "non-linear place-time model" of Peter Assmann and Quentin Summerfield (1990) produced "the most accurate estimates of the f_0's of pairs of concurrent synthetic vowels and comes closest to predicting the identification responses of listeners to such stimuli" (680). A combined temporal-spectral approach may enable the brain to separate voices in everyday situations of extraordinary complexity (the cocktail party effect; Bregman 1990) as well as multivoiced music (Bello et al. 2006). When Aniruddh Patel and Evan Balaban (2001) aimed to investigate the relationship between pitch and the "timing of stimulus-related cortical activity" (839), they commented that "because of frequency-dependent time delays in the cochlear, either [spectral or temporal] method of extraction (or a mixed method) could contribute to cortical activity in which the timing varies with the perceived pitch of the stimulus. These findings also do not rule out subtle spatial changes in the magnitude of cortical activity during pitch perception, because these may be reflected in aspects of the brain signal that our study was not designed to detect" (843). In general, then, pitch can correspond to (or somehow emerge from) many different aspects of brain activity, of which any given empirical study has access to only a part.

In this regard, the term "spectral pitch" can be misleading. Spectral pitches do not depend only on the spectrum or frequency-domain representation. Like virtual pitches, they also depend on the temporal pattern of electrical responses in the auditory nerve (B. C. J. Moore 1989). It is also misleading to call Terhardt's model "spectral." A sound spectrum is part of World 1, but his model focuses on the World-2 relationship between spectral and virtual pitch. The model is only spectral insofar as the calculations are carried out in the pitch domain for convenience, assuming a one-to-one correspondence between the pitch and the frequency of a pure tone of constant SPL.

Brian Moore (1989, 1990) presented a model of pitch perception that incorporated both spectral (tonotopic) and temporal (timing) information. Like Terhardt's, it began with acoustic input to a filter bank: each hair cell on the basilar membrane may be regarded as an auditory filter that responds to a limited range of frequencies. Spectral pitch was assumed to depend on a mixture of spectral information (position on the basilar membrane) and temporal information (in auditory nerve fibres). The next stage involved the recognition of periodic or harmonic patterns in an approach that differed from Terhardt's: "In [Terhardt's] model each partial gives rise to potential pitches corresponding to subharmonics of that partial, and the perceived pitch is that for which the number of coincidences is greatest. In our model these subharmonic pitches correspond to the interspike intervals of 1, 2, 3, 4, 5 . . . ms, and these correspond to pitches of 1000, 500, 333.3, 250, 200 . . . Hz" (B. C. J. Moore 1989, 186). Note the mathematical equivalence of periodicity and harmonicity in this case.

Terhardt's Algorithm

In this book, I am following the approach of Terhardt et al. (1982b), for the following reasons:

- It assumes that the relationship between spectral and virtual pitches is partly learned and can be affected by musical experience (Preisler 1993). For example, gamelan and Western musicians hear different pitches in nonharmonic gamelan sounds (McLachlan, Marco, and Wilson 2013; Parncutt 1989).
- It acknowledges the existence of pitch shifts and tries to model them. It assumes plausibly that virtual-pitch shifts can be predicted from spectral-pitch shifts (due to changes in SPL and masking) and learned harmonic patterns (slightly shifted or distorted spectra of harmonic complex tones).
- Its predictions are limited to aurally relevant aspects of the sound signal. Unresolved partials are included in the calculation insofar as they are aurally relevant.

For an accessible description of how Terhardt's algorithm deals with different combinations of frequencies in an input spectrum and how it applies Bregman's principle of harmonicity in different cases, see Huron (2016, 27–32). The algorithm is based on simple, widely accepted assumptions:

- Only aurally relevant partials influence pitch. Spectrally unresolved upper harmonics may or may not be aurally relevant.
- Spectral pitch salience is a measure of a partial's aural relevance. It depends on SPL relative to masked threshold and frequency relative to the spectral dominance region, the ear being more sensitive to spectral pitch in a log-frequency region centered on about 700 Hz.
- The greater the number and salience of spectral pitches matching a harmonic template, the more salient is the virtual pitch at or near the fundamental frequency.
- Aurally relevant partials that correspond to lower harmonics of a virtual pitch have a greater effect on its salience than those corresponding to higher harmonics.
- Partials can be mistuned relative to the harmonic template by up to about a semitone (6 percent) and still contribute to virtual pitch at the template's fundamental; mistuning gradually reduces the contribution.

The small number of assumptions makes the model easy to falsify, and the lack of references to auditory physiology makes the model less arbitrary.

Terhardt's algorithm was in part inspired by Helmholtz. Here is an extract from Helmholtz's (1856) paper on combination tones: "Now, the ear is accustomed to hearing the musical tones of musical instruments, the human voice, etc., *always in the same composition repeatedly*, so that they become to the ear specific and familiar sense perceptions, about whose composition it has no occasion to reflect, indeed little more than we habitually make clear to ourselves that the sensory intuition of a physically extended object is composed of two different retinal images in both eyes. Indeed, in the everyday practice of our senses, we only attend to our sense perceptions at all to the extent that they enable us to recognize objects and events of the outside world" (Steege 2012, 73; italics in original).

The composition of musical instrument tones and voice sounds to which Helmholtz was referring is the harmonic series formed by the audible partials, which according to Terhardt becomes familiar to the auditory system in early life. That familiarity enables the ear to track the fundamental frequency and distinguish different timbres. But we are not aware of these recognition processes. Instead, we are aware of the sound source as constructed by our perceptual system. Philip Glotzbach and Harry Heft (1982) described how our perception constructs the current state of the outside world:

> Sensory input remains ambiguous, or at least incomplete. As a result, the organism must construct hypotheses to deal with it. (109)
>
> [In] traditional perceptual theories . . . animals are forced to gamble their well-being by constructing perceptual hypotheses of the "most likely" distal environmental conditions based on incomplete proximal cues. But one can ask: if the environmental regularities are available, why could not species have evolved the capacity to detect them directly, without having to resort to mediating probabilistic constructs? (110)

Our sensory systems are constantly constructing a model of the outside world based on available real-time sensory inputs and previous experience of environmental objects and the stimuli they usually produce. That principle applies to any aspect of perception in any sensory modality. The construction happens spontaneously and effortlessly, and indeed without awareness that construction is happening. Instead, we are aware of the output of the process: the world as constructed by our senses. In vision, for example, we are not aware of differences between the images picked up by the left and the right eye. Instead, we are aware of the distance of an object from us and the object's 3D features, which the distance between the eyes enables us to perceive.

One might argue that Helmholtz was implying (or even predicting) an ecological approach to perception, in which it is primarily the interaction between the perceiver and the environment that determines what we attend to. But his approach also anticipated today's cognitive approaches by dividing perception into two separate unconscious stages: sensation and inference (Barlow 1990). He also correctly implied that there is no contradiction between ecological and cognitive approaches.

Subjectivity of Pitch

Skeptical monists may doubt the existence or relevance of virtual pitches. Here are three possible reasons:

- If only the physical World 1 exists, virtual pitches do not exist by definition, although they may somehow exist in the brain as emergent properties of neural networks that cannot yet be reliably predicted.

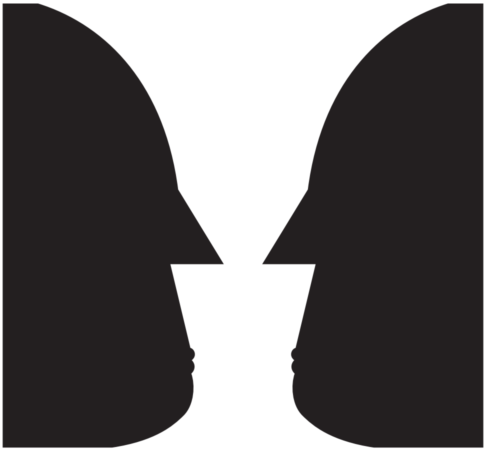

Figure 13.3
The Rubin vase. The figure can be seen as either a white vase or two black faces. (From Wikipedia: Rubin vase; Gestalt psychology.)

- Virtual pitches tend to “disappear” (become inaudible) when the listener focuses attention on them, just as holistic aspects of a pattern tend to disappear when we focus our attention on an element. In vision, the boundaries of the ghost triangle in the Kanizsa illusion and the dots at intersections in the Hermann grid illusion may disappear when we attend to them; different physiological bases have been proposed (Schiller and Carvey 2005). In the sound of a church bell, the first pitch we hear is the strike tone, which is a virtual pitch that depends on recognition of a harmonic pattern among the bell’s audible partials. If we then focus our attention on the hum tone (the most salient spectral pitch), the virtual pitch becomes less salient—comparable with switching attention between the white Rubin vase and the black pair of faces in figure 13.3. Readers are invited to listen to an Internet recording of Vienna’s biggest church bell, the *Pummerin*, whose strike tone is about F3 and whose hum tone is about D♭3. As the sound dies away, the (virtual) F3 becomes less prominent, and the (spectral) D♭3 becomes more so.
- There are big individual differences in pitch perception. People tend to be either fundamental listeners or spectral listeners (Seither-Preisler et al. 2007). Presented with harmonic complex tones with missing fundamentals, overtone listeners are more likely to hear spectral pitches (the pitches

of upper partials) consciously, whereas fundamental listeners are more likely to hear the pitch at the missing fundamental. Spectral listeners may be more likely to doubt the existence of virtual pitches.

Despite these uncertainties, there are good reasons to believe that virtual pitches exist. Beyond the empirical and theoretical evidence reported in this book, there is everyday experience of speech: we effortlessly perceive fundamental frequency trajectories, even when speech is masked by background noise (conversation, ocean waves, road traffic, music; J. H. Song et al. 2011). That is a practical consideration, and there is also a philosophical one: if most pitches heard in speech and music were virtual, it would be surprising if they did not exist. By the same token, if all of humanity was engaging with certain nonessential behaviors, calling them "music," and explaining their value in terms of subjective experience, it would be surprising if that experience did not exist.

The standard experiment for measuring pitch (described above) is objective in the sense that the results cannot be biased by the experimenter. A computer plays sound over headphones and automatically shuffles the order of trials for each new participant. It makes no difference if the experimenter accidentally smiles when listeners should reply "yes" and frowns when they should reply "no." If the results show that listeners consistently perceive a pitch, then it exists.

Pitch is inherently subjective, in two different ways. First, there are individual differences in the salience and shift of individual pitches. Second, pitch is part of Popper's World 2. To measure pitch, a human listener has to adjust the frequency of a pure tone until its pitch matches that of a test sound. That process relies on consciousness and attention, and hence on subjectivity.

To understand the inherent subjectivity of pitch better, compare it with another inherently subjective entity: timbre or sound quality. Skipping over attempts to define or operationalize timbre in the literature for the moment, timbre might loosely be defined as "what a sound sounds like" by comparison to known models in different sensory modalities (e.g., a bright light, a rough stone, chocolate, or a clarinet). Like pitch, timbre depends on various physical properties of a tone—not only spectral envelope, as is often assumed. Unlike pitch, timbre cannot be measured on a one-dimensional scale, but it does have perceptual dimensions that can be measured; they may correspond physically to spectral energy distribution or spectral fluctuation (Grey 1977). Those are good reasons to regard timbre as quite a different thing from the temporal and spectral envelopes upon which it depends—just as there are good reasons to regard pitch as different from frequency.

We can also regard a perceived tone as a subjective entity, of which pitch as an attribute. In both cases, a pitch is a *quale*—a unit or instance of subjective, conscious experience. The quale in question can be seen either as a property of a physical phenomenon (the tone—an oscillation in air pressure) or as a property of another quale (the tone sensation—a subjective entity).

Whereas some qualia are ineffable, pitch (as defined) can be reliably measured—contradicting claims in the literature that qualia cannot be measured. A standard psychoacoustic experiment allows us to measure pitch reliably. The result is usually presented as an average across several listeners on the assumption that they are experiencing almost the same thing. I will assume that these global claims are true regardless of how the term "qualia" is understood in detail, or what detailed questions about (musical) qualia are of concern to philosophers (e.g., Schiavio and van der Schyff 2016).

The discipline of psychoacoustics (more generally, psychophysics) can be seen as a form of quantitative phenomenology. Subjective experiences are measured, and results from different participants are averaged to improve reliability. The approach is comparable with what Daniel Dennett (2003) called heterophenomenology or a "third-person approach to human consciousness" (19)—phenomenology of another, not of oneself. Dennett was interested in qualitative description, whereas we are here dealing with quantitative measures. The advantage of Dennett's idea, like the critical phenomenology of Max Velmans (2007), is that it enables a more objective description of subjectivity than mere introspection. The disadvantage of heterophenomenology is that ineffable aspects of experience are excluded.

This discussion implies that reliable information about subjective entities is possible, which in turn justifies believing in the existence of those entities. The reliable information in question includes replicable findings of psychoacoustic experiments if we ignore the uncertainty in the measurements; after all, physical measurements are also generally uncertain.

To clarify the inherent subjectivity of pitch further, imagine an experiment on pitch perception in monkeys. In each trial, the monkeys hear two successive tones. The pitch of the second tone is either higher or lower than the first. The monkeys are trained by operant conditioning to respond differently to rising and falling tone pairs. In experiments with New World monkeys, Xindong Song and colleagues (2016) reported that "the subject had to detect the F0 change and respond by licking at a feeding tube placed in front of its mouth during the response window ("hit") to receive a food reward" (785). The research team found that the relationship between spectral and virtual pitch in monkeys is much the same as it is in humans. So far, so good.

Now imagine a new experiment to test the perception of pitch shifts or pitch ambiguity in monkeys. Pitch shifts are small changes in pitch due to physical intensity or masking (typically, less than a semitone), whereas pitch ambiguity means pitch can flip across larger intervals (e.g., an octave). Monkeys and humans hear complex tones in a similarly holistic manner. But in the case of monkeys, we have no access to their subjective impression of pitch because we have no access to their private world. As in humans, their impression could be affected by diverse aspects of neural processing that correspond in complex ways to both time-domain and frequency-domain aspects of sound. Proponents of 4E cognition would agree that a monkey's impression of pitch also depends on interactions between the brain and the rest of the body or between the body and the environment.

Like humans, monkeys have enormous numbers of neurons in their brains, and the number of synapses joining them together is greater again. Trying to understand how that complexity relates to pitch judgments of monkeys is like trying to solve the mind-body problem in humans. This comparison between humans and monkeys is valid regardless of whether monkeys have consciousness and, if so, of what kind.

The point applies equally for both spectral and virtual pitch. Just because spectral pitch corresponds to an approximate position on the basilar membrane and/or to information contained in neural timing patterns does not mean that spectral pitch is a physical parameter or somehow more physical than virtual pitch. Like virtual pitch, spectral pitch can only be determined by an experimental procedure that involves subjective appraisal and comparison by a conscious observer. Like virtual pitch, spectral pitch is purely subjective in nature.

Neuroscience of Pitch

The physical states and processes of the human brain that correspond to experiences of pitch are complex and difficult to investigate (Plack et al. 2014). Deborah Hall and Christopher Plack (2009) found evidence for neural processing corresponding to pitch in the lateral Heschl's gyrus, planum temporale, temporo-parieto-occipital junction, and prefrontal cortex. For perceiving and producing pitch in speech, Martin Meyer and colleagues (2004) discussed the roles of the superior temporal gyrus, dorsal striatum (putamen and caudate nucleus), right Rolandic operculum (premotor cortex), left inferior frontal gyrus (Broca's area), left inferior precentral sulcus, and left Rolandic fissure.

Different brain areas can contribute to pitch perception in different ways, depending on the linguistic, musical, or environmental function of pitch (cf. Gandour et al. 1998). There are also individual differences: different brains can process pitch in different ways. That being the case, we might pose an interesting armchair question: Imagine that we knew everything that there was to be known about the physical state of a given brain, including the state of every neuron and every synapse, from one moment to the next. Would we be able to predict perceived pitch from physical measurements? Given the inherent complexity of the mind–brain relationship, the answer is unclear.

I prefer not to jump to conclusions about such a difficult problem, and I agree with Terhardt that a music-theoretically interesting model of pitch perception should follow cybernetic principles and not depend on neurophysiology. It is possible that a future neuroscientist will achieve a breakthrough that will make the preceding arguments obsolete. That would be great, but in the meantime, we should remain scientifically cautious.

Temporally based models of pitch focus on the phase locking of neural firing along auditory pathways. They assume that pitch is related to the rate of neural firing because the two are clearly correlated, at least up to a limit of a few thousand hertz (Verschooten et al. 2019). Research into these issues often employs artificial or unnatural sounds, such as complex tones in which the phase relations between the partials are independently manipulated and all variables are carefully controlled. But we can also learn about pitch perception in experiments using real, ecologically valid environmental sounds in all of their complexity. We need to take a broad perspective that includes many different empirical observations. Plack and Oxenham (2005) pointed out that "when developing a model of pitch perception, or when identifying a cell type or brain region that may be involved in pitch perception, it is important to ensure that the results are consistent with the wide range of psychophysical observations, and not to focus on a single property of pitch that may provide an easy solution" (7).

Physiologically and temporally based models of pitch perception involve different physical processes, including autocorrelation in the timing of neural impulses in the auditory nerve (Cariani 2001), cancelation (Cheveigné 1998), and cochlear filtering and coincidence detection (Shamma and Klein, 2000). Neural pulse codes may be channel-based, temporal-pattern, or time-of-arrival codes (Cariani 1999). It is tempting to explain aspects of musical harmony on the basis of such models (e.g., Cariani 2001; Tramo et al. 2001), but the approach is misleading if it assumes Just tuning and intonation (ratios of small integers; see chapter 7). Mark Tramo and colleagues (2001) noted that “for the consonant intervals (the fifth and fourth), the pattern of major and minor peaks in the autocorrelation is perfectly periodic, with a period related to the fundamental bass . . . suggest[ing] that the consonance of harmonic intervals reflects regularities in their temporal fine structure” (100–101). Apart from assuming that musical intonation is ideally Just, this claim skims over three kinds of complexity: that of music (which includes both deliberate and accidental variations in intonation), that of the brain (which comprises enormous numbers of neurons that could potentially be involved in a single perceptual act such as perceiving a musical interval), and that of relevant philosophy and metaphysics (in which the mind-body problem remains unsolved).

Pitch shifts and octave stretch pose additional problems. Perceived pitches often depart significantly from pitches predicted on the basis of neural-temporal patterns. There is also research showing that tonotopic representations (frequency–place mappings) are essential for pitch perception (Oxenham et al. 2004). That does not mean the neural-temporal patterns are unrelated to pitch. On the contrary, it suggests that pitch perception generally depends on a complex mixture of, and possibly interaction between, temporal and spectral (or tonotopic) factors.

Creators of temporal models of musical interval perception are in good company. Beliefs in music of the spheres and musical intervals as number ratios have been seriously held by scientists as eminent as Nicolaus Copernicus in the fifteenth century; Leonardo da Vinci and Gioseffo Zarlino in the sixteenth; Galileo Galilei, René Descartes, Johannes Kepler, and Marin Mersenne in the seventeenth; Isaac Newton, Leonhard Euler, Jean-Philippe Rameau, and Immanuel Kant in the eighteenth; and even (to a lesser extent) Hermann von Helmholtz in the nineteenth (Pesic 2014; Proust 2009). Besides, neural periodicities can indeed correspond to pitch. When a tone is perceived, temporal behaviors of neurons are often phase locked to stimulus waveforms along auditory pathways (Cariani 1999; Langner 1997), and their periods correspond inversely to perceived pitch. But there is a danger of jumping to conclusions. Correlation does not imply causality. Just because something that we experience is correlated with something happening in the physical world does not mean that the physical states or changes are causing the experience.

Misunderstandings

Differences between schools of thought in psychoacoustics or pitch perception may involve simple operational definitions of pitch or audibility. I am assuming that pitch is measured by the frequency of a pure tone that is judged subjectively to have the same pitch as a successively presented test sound. That is true regardless of what spectral or temporal aspect of the signal is causing the pitch. For example, if pitch is perceived in a fast train of clicks, the pitch is still measured by the same psychoacoustic procedure. I am also assuming that a partial within a complex sound is audible if a listener notices a change when the partial is turned on and off. Others may define pitch and audibility differently.

Misunderstandings can also involve differences in unstated assumptions about the nature of reality (metaphysics). The auditory scene is often treated as if it were physical or physiological in nature; I have argued that it is not. Of course, the auditory scene has a physical or physiological foundation, but the measurements upon which it is based are psychoacoustic—mediated by human consciousness. We should not pretend that we are making physical measurements when in fact we are not.

For the music-theoretical purposes of this book, I am sidestepping the physiological complexity of complex tone perception and corresponding unresolved empirical and philosophical issues, and focusing instead on the psychoacoustic relationship between physical parameters on the one hand (the frequencies, amplitudes, and phase relations of the partials of a complex sound in air) and experiential parameters on the other (pitch, loudness, timbre, perceived duration). This psychoacoustic approach is appropriate for investigating the physical and psychological foundations of MmT as an experiential, subjective, and cultural phenomenon, and for addressing questions such as: Why is (Western) musical structure based on hierarchically structured pitch-time patterns? Why is conventional music notation, like representations of the auditory scene in psychoacoustics, similar to a graph of pitch against time? Both these questions refer to the ear's extraordinary sensitivity to pitch by comparison to loudness, timbre, or phase (see chapter 6).

What we know about the auditory scene (what it contains, what it does not, what parts of it attract attention, and so on) comes from psychoacoustic experiments. Data of that kind are always mediated by the consciousness of the listener. In that sense, the auditory scene is a fundamentally subjective entity. It is a representation of subjective experience. While it could not exist without the underlying physiology, the underlying physiology is not what is measured in experiments of that kind. If a measurement is based on a subjective judgment, which is mediated by consciousness, it should be labeled "subjective"—if only for scientific clarity and consistency.

In the three-worlds concept of Karl Popper and John Eccles (1977), the pitches that we perceive in the auditory scene belong to World 2 (subjective)—not World 1 (physical). The pitches that we perceive certainly *depend on* World 1, but since our perception is inherently subjective, we should not jump to conclusions about that dependency. We need to keep our feet on solid scientific ground.

Many psychologists, including many music psychologists, disagree with the previous arguments. The past century has seen a tendency toward physicalism—the belief that only the physical world exists, denying the existence of subjectivity (Papineau 2013). Within that broader academic and social context, many colleagues in music psychology are studying aspects of musical subjectivity, and some are devoting their lives to that task. Of them, a large proportion may paradoxically believe that the object of their research does not exist. They are right, from a physical perspective, that only the physical world exists. They are also right that most of the sciences operate most successfully on a physicalist-monist assumption. The trouble is that they are studying art, and art is inherently subjective. Without subjectivity, music would be meaningless.

This book is about the pitch structures that people experience when listening to music. For every such experience, highly complex physical correlates in the brain can be identified. These correlates can and should be studied in detail in order to understand the anatomic structures and their biological functions better, both within and between neurons. But the observations of neuroscientists are limited to the physical world. Regardless of unpredictable future technological developments, and despite the fascinating fantasies of science fiction, neuroscientists will never directly observe the subjective experience of live subjects within their brains.

Psychoacoustic Foundations of Major-Minor Tonality

By: Richard Parncutt

Citation:

Psychoacoustic Foundations of Major-Minor Tonality
By: Richard Parncutt
DOI: 10.7551/mitpress/15050.001.0001
ISBN (electronic): 9780262377362
Publisher: The MIT Press
Published: 2024

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data
Names: Parncutt, Richard, 1957– author.
Title: Psychoacoustic foundations of major-minor tonality / Richard Parncutt.
Description: Cambridge, Massachusetts : The MIT Press, 2024. | Includes bibliographical references and index.
Identifiers: LCCN 2023007827 (print) | LCCN 2023007828 (ebook) | ISBN 9780262547352 (paperback) | ISBN 9780262377379 (epub) | ISBN 9780262377362 (pdf)
Subjects: LCSH: Tonality. | Psychoacoustics. | Chords (Music) | Musical intervals and scales.
Classification: LCC ML3811 .P37 2024 (print) | LCC ML3811 (ebook) | DDC 781.2/6—dc23/eng/20230606
LC record available at https://lccn.loc.gov/2023007827
LC ebook record available at https://lccn.loc.gov/2023007828