CHAPTER 8

Pitch Perception

Adrianus J. M. Houtsma

I. INTRODUCTION

Pitch is defined by the American National Standards Institute (1973) as "that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from high to low." Pitch is a particularly important attribute of sound. It is an essential element for features such as melody and harmony in music, and it conveys the bulk of the prosodic information in speech. Like loudness and timbre, it is a subjective attribute that cannot be expressed in physical units or measured by physical means. In the case of a *pure tone*, its primary objective correlate is the physical attribute frequency, but the tone's intensity, duration, and temporal envelope also have a well-established influence on its pitch. If a tone is *complex* and contains many sinusoids with different frequencies, which is usually the case with natural sounds, we may hear a single pitch as, for instance, in the case of a single note played by a clarinet. We may also hear a cluster of pitches as, for instance, a chord being played by a group of instruments. We may even hear individual partials as sinusoids, all having their own pitches. Even sounds that are not formed of well-defined discrete partials can evoke pitch sensations. This will be referred to as *nontonal* pitch.

Our auditory memory seems to be particularly good at storing and re-

Hearing

trieving pitch relationships, given that most people can easily recognize tunes or melodies and sing them more or less correctly. This ability to recognize and reproduce frequency ratios is often referred to as *perfect relative pitch*. Some people possess the ability to identify the pitch of sounds on an absolute, nominal scale without any explicit external reference. This relatively rare ability is referred to as *perfect absolute pitch*.

In this chapter we will first discuss the sensation of pitch evoked by pure tones, its dependence on various physical attributes of the signal, and our sensitivity to changes in frequency. We will consider complex tones and show how a single holistic pitch percept is determined by fundamental frequency as well as harmonic partials. A third class of sounds to be discussed consists of signals having continuous spectra, with a temporal or spectral regularity or with a spectral discontinuity. Such sounds can evoke pitch sensations corresponding with modulation frequency, spectral ripple, or edge frequency. We will discuss how pitch is internally represented, either as part of a musical scale or an intonation contour in speech, or in isolation, as in the case of absolute pitch. Finally, the multidimensional nature of pitch will be discussed in terms of the attributes pitch chroma, tone height, and harmonic proximity.

II. PURE TONES

A. The Mel Scale

There are various methods for measuring how the pitch of a pure tone depends on its frequency. One can obtain a pitch–frequency relation by magnitude estimation. One can also use a "doubling" or "halving" method in which subjects adjust the frequency of a comparison tone until it subjectively sounds twice or half as high as the pitch of a test tone with a frequency set by the experimenter. The classical result of such experiments is the *mel scale* measured by Stevens, Volkmann, and Newman (1937) and shown in Figure 1. The scale, obtained by the method of pitch halving, has an arbitrary pitch reference of 1000 mels at a frequency of 1000 Hz. A tone that sounds, on average, twice as high receives a value of 2000 mels, whereas a tone that sounds only half as high has a pitch of 500 mels. One can clearly see that pitch, expressed in mels (the unit is derived from *mel*ody), is not identical to frequency and not even linear in frequency. There is a direct and simple relationship between the mel scale of pitch and the critical-band scale (bark scale) for frequency resolution in the ear as discussed in Chapter 5. Zwicker and Feldtkeller (1967) pointed out that 1 bark is exactly 100 mels, which implies that the scales are essentially the same. This is because pitch, as measured by Stevens et al., is apparently determined by the center of excitation activity along the basilar membrane, which is also reflected in the

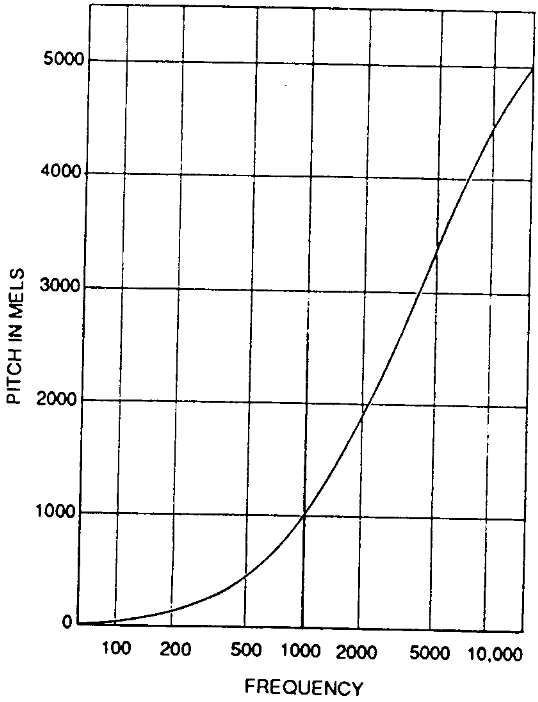

FIGURE 1 The relation of pitch (in mels) to the frequency of a pure tone. A 1000 Hz tone is arbitrarily assigned a value of 1000 mels. (From Stevens et al., 1937, reprinted with permission.)

critical-band or bark scale. Although the mel scale is based on empirical and scientific results, musicians may find it difficult to reconcile such a scale with the familiar subjective musical intervals of fifths, octaves, or semitones that they tend to use as relative scale units. It is probably for that reason that the mel scale never became quite as popular as the comparable *sone scale* for loudness.

B. Dependence on Intensity

Although the mel scale suggests that the pitch of a pure tone is simply determined by its frequency, the perceived pitch also depends on some other factors, one being intensity. If one measures for a group of subjects how, on average, the pitch of a pure tone changes with the tone's intensity, one typically finds that (1) for tones below 1000 Hz the pitch decreases with increasing intensity, (2) for tones between 1000 and 2000 Hz the pitch remains rather constant, and (3) for tones above 2000 Hz the pitch tends to rise with increasing intensity. Stevens (1935) reported the first data on this phenomenon, coming mostly from one listener, which are shown in Figure 2. Subsequent investigations have shown that for most people the magnitude of the pitch shift effect is smaller than was reported by Stevens (Verschuure & van Meeteren, 1975) and that the effect varies considerably between individual listeners (Morgan, Garner, & Galambos, 1951; Terhardt, 1974a). Interquartile ranges found by Morgan et al. (1951) have been superimposed on Stevens' data in Figure 2. For very short tone bursts, less than 40 ms, an increase in intensity always seems to lower the pitch, regardless of the tone's frequency (Rossing & Houtsma, 1986). This is probably also the reason why the pitch of such a very short tone burst depends on the shape of its temporal envelope, with the lowest pitch always being obtained with a constant-amplitude on–off gate function (Hartmann, 1978; Rossing, & Houtsma, 1986).

C. Influence of Partial Masking

The simultaneous presence of other tones or noise may also alter the perceived pitch of a pure tone. If the interfering tone or noise band is just below the frequency of a test tone, the pitch of this test tone is always increased, sometimes by as much as a semitone (Terhardt & Fastl, 1971), as is illustrated in Figure 3. Interfering sounds above the test tone frequency have a much less consistent effect.

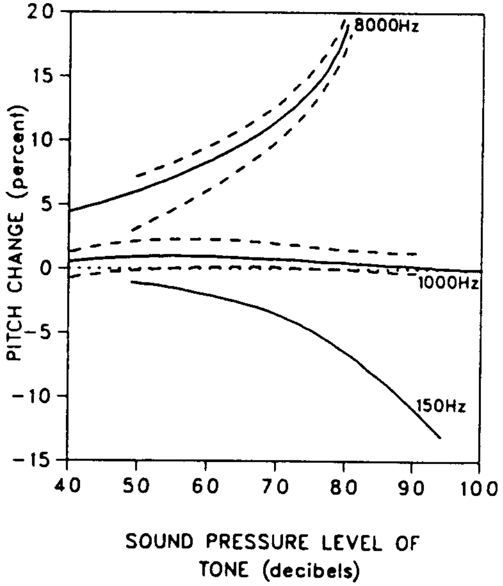

FIGURE 2 Pitch change as a function of sound pressure level of a pure tone. Solid curves: mean data from Stevens (1935). Dashed curves: 25th (lower) and 75th (upper) percentile of distribution of pitch changes in 18 ears, measured by Morgan et al. (1951).

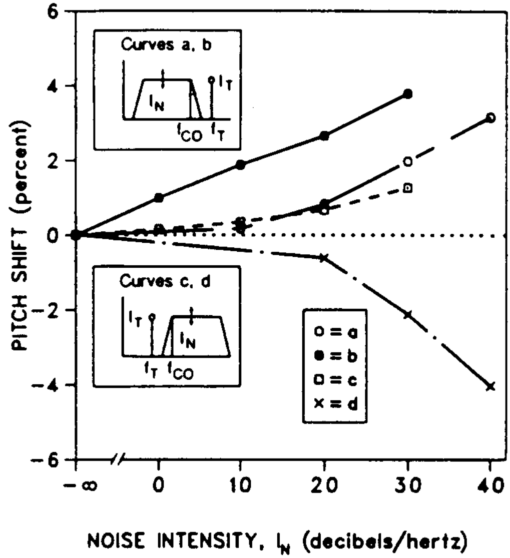

FIGURE 3 Pitch shift of a sinusoidal test tone induced by bandpass noise just below (a and b) and just above (c and d) the test tone frequency. Sound pressure levels of test tones were 50 dB, and frequencies were (a) 300 Hz, (b) 3800 Hz, (c) 3400 Hz, and (d) 100 Hz. (After Terhardt & Fastl, 1971.)

D. Binaural Diplacusis

Finally, the pitch sensation of a pure tone also typically depends somewhat on the ear to which it is presented. If a subject is asked to adjust the frequency of a comparison tone in one ear so that it matches the pitch of a test tone in the other ear, the frequencies will often come out slightly but consistently different. This effect, which is found to some extent in every listener, is known as *binaural diplacusis*. Interaural pitch differences are normally less than 2%, and the interaural frequency difference function may change slowly with time, as has been measured by van den Brink (1970) and is shown in Figure 4.

E. Frequency Discrimination

If two sinusoidal tones of different frequency are presented sequentially, there is some smallest frequency difference below which listeners can no longer tell consistently which of the tones is higher. A frequency difference resulting in 75% correct responses in a two-interval, two-alternative forced-choice paradigm is usually referred to as a *just noticeable difference,* or jnd. Classical frequency jnd data were measured by Shower and Biddulph (1931) with a method of frequency modulation detection. Modern data were provided by Moore (1973), who measured frequency jnds as a function of tone

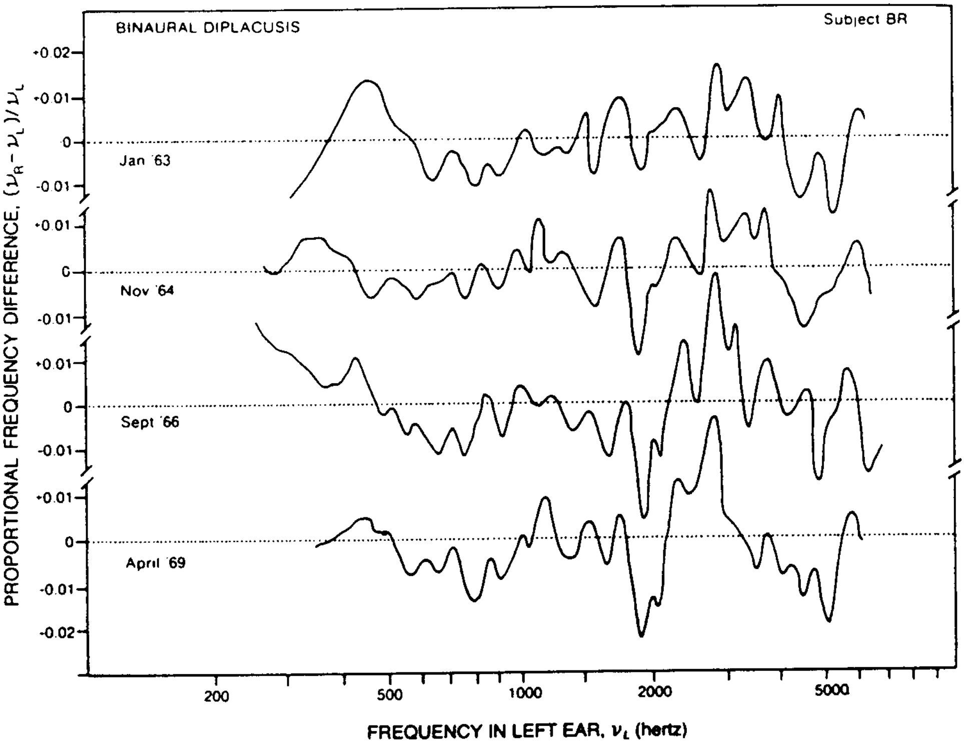

FIGURE 4 Binaural diplacusis patterns of one subject measured at intervals of several years. (From Brink, 1970, reprinted with permission.)

duration, and by Wier, Jesteadt, and Green (1977), who measured jnds as a function of tone intensity. A summary of Moore's data is shown in Figure 5, and of the data by Wier et al. in Figure 6.

The data of Figure 5 and 6 clearly show that the accuracy of our hearing system for distinguishing sequential tones of different frequency is much greater than the ability to resolve these tones (see Chapter 5). The large difference between the 0.1–0.2% frequency discrimination threshold and the approximately 10% frequency separation required to resolve simultaneous tones has sometimes been presented as a paradox and has been a reason for assuming the presence of "neural sharpening" mechanisms in the central auditory system. The reader should realize, however, that frequency discrimination behavior and frequency resolution in the auditory periphery have, in principle, very little to do with one another. Discrimination limits are imposed primarily by the amount of noise in the system. If there were no noise, one would be able to discriminate tones with an arbitrarily small frequency difference, no matter how steep or shallow were the slopes of peripheral auditory filters. Given that the frequency encoding process in the

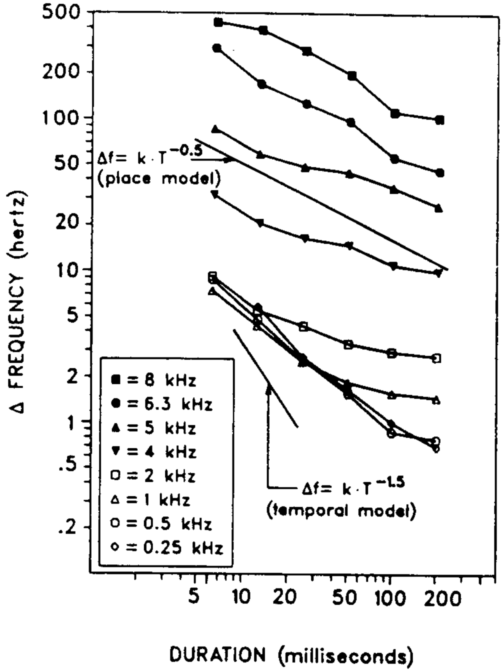

FIGURE 5 Just noticeable frequency differences as a function of stimulus duration. Sinusoidal tones had a constant loudness level of 60 phons. Line segments represent predictions of Siebert's (1970) place model (top, slope = −0.5) and temporal model (bottom, slope = −1.5). (Data from Moore, 1973.)

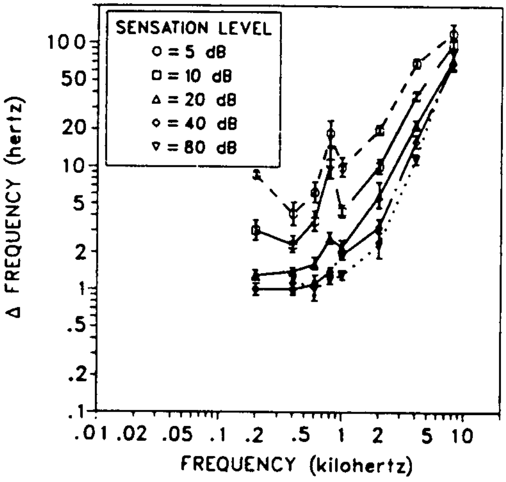

FIGURE 6 Just noticeable frequency differences as a function of frequency, with sensation level as parameter. Tone durations were 500 ms. (Data from Wier et al., 1977, reprinted with permission.)

auditory system is noisy, however, the resolution power at the periphery will show up as a model parameter in any stochastic frequency-coding model. For instance, Siebert (1970) has shown that optimal use of neural firing rate information across fibers of the auditory nerve, assuming cochlear filtering in accordance with the classical observations by von Békésy, predicts a frequency discrimination performance comparable to the data of Shower and Biddulph (1931). Performance is predicted to be proportional to the inverse square root of tone duration. Optimal use of all temporal information in firing patterns yields a predicted performance that is much better than is observed under any condition and is proportional to duration raised to the power −1.5. One can see from the time-dependence slopes shown in Figure 5 that high frequencies tend toward predictions of the place model, whereas the lowest frequencies (250 and 500 Hz) show a duration dependence that is more in accordance with a time model.

III. COMPLEX TONES

A. Historical Background

Between 1840 and 1850 an interesting discussion took place in the *Annalen für Physik und Chemie* between Ohm and Seebeck about the pitch of a complex tone. Such a tone is composed of several sinusoidal tones, the lowest in frequency being the *fundamental*, and the others (*harmonics*) having frequencies that are multiples of the frequency of the fundamental (see also Chapter 1). Seebeck (1841) presented observations on sounds made with a mechanical siren. These sounds were periodic, containing controllably suppressed odd harmonics. Seebeck described how the pitch he associated with the sound as a whole always seemed to follow the fundamental, even if this fundamental component was very weak. He concluded that the fundamental frequency is not the only determinant of pitch, but that the upper harmonics also contribute to the subjective pitch sensation. Ohm (1843) argued that our ears perform a real-time frequency analysis similar to the mathematical formulation of Fourier, where the frequency of the lowest spectral component determines the pitch of the complex, and the other frequencies determine the sound's timbre. The strong fundamental pitch sensation in the absence of acoustic power reported by Seebeck therefore had to be based on an illusion. Twenty years later Helmholtz (1863) chose the side of Ohm in this debate and thereby settled the issue for almost a century to follow.

Just before the Second World War Schouten (1938) rekindled the Ohm–Seebeck debate by demonstrating that Seebeck's conclusion was basically correct. With his optical equipment he could generate periodic complex tones devoid of any acoustical power at the fundamental frequency. Schouten was able to show that the pitch sensation associated with the *missing* *fundamental*, as it later became known, could not be explained as a nonlinear difference tone generated at the auditory periphery, as first Helmholtz (1863) and later Fletcher (1924) had argued. According to Schouten, the pitch sensation is caused by neural detection of periodic fluctuations in the envelope pattern of clusters of harmonics that the ear fails to resolve. If spectral resolution is insufficient, two or more summed harmonics will appear at the output of the cochlear filter. The periodicity of the envelope of such a summed signal is the same as the periodicity of the fundamental, even if the fundamental is physically absent. It can be picked up through phase locking by fibers of the auditory nerve and transmitted to central parts of the brain. Since insufficient cochlear resolution is an essential element of Schouten's pitch theory, this theory became known as the *residue theory of pitch* (Schouten, 1940).

Soon it became clear, however, that Schouten's residue theory also failed to provide an adequate explanation of new experimental findings. Ritsma (1962) found a clear upper limit to the harmonic order beyond which no tonal residue, that is, pitch, is heard. He also found that the existence region for the tonal residue extends to combinations of harmonics that the cochlea should be able to resolve, which is in contradiction with the essence of the residue theory. Some years later Ritsma (1967) and Plomp (1967) found that the best harmonics to convey a pitch sensation of a missing fundamental are on the order of 3, 4, and 5. In this so-called dominant region, harmonic frequencies differ by 25% or more and should, as has been discussed in Chapter 5, be well resolved in the periphery of the auditory system. Perhaps the most direct evidence against the residue theory was the finding by Houtsma and Goldstein (1972) that two successive simultaneous harmonics with frequencies nf_0 and $(n + 1) f_0$, presented to different ears, evoke an equally effective fundamental pitch percept as a monotic or diotic presentation of the same two harmonics. In the dichotic case, with each harmonic going to a different ear, there is no physical interference or cochlear residue. The experimental results force one to conclude that the pitch of complex tones is mediated primarily by a central mechanism that operates on neural signals derived from those stimulus harmonics spectrally resolved in the cochlea. Modern pitch theories therefore almost always contain elements of *central* pitch processing.

B. Template Theories of Pitch

In this class of theories it is assumed that, at some central processing stage in the brain, a spectral template is matched to frequencies or frequency transformations of those stimulus partials that are resolved in the cochlea. One of these pitch theories is the optimum processor theory of Goldstein (1973), schematically illustrated in Figure 7. The theory assumes that the frequen-

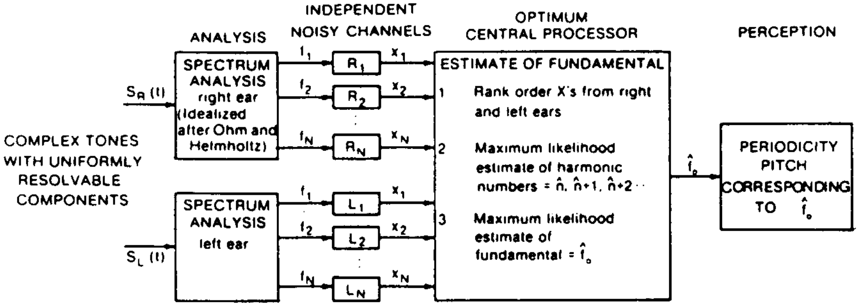

FIGURE 7 Schematic representation of the optimum processor model of Goldstein (1973). (Reprinted with permission.)

cies f_i of spectrally resolved stimulus partials are transformed into Gaussian random variables x_i, with means equal to f_i and variances that are functions of f_i only. All amplitude and phase information is ignored. A central processor assumes that the input numbers x_i are noisy representations of harmonic frequencies and makes an optimal estimate of the unknown harmonic numbers and fundamental frequency. The variance function is the only free parameter of the model. This function represents the noise in the frequency coding process in our auditory system. It causes the central processor to sometimes make an incorrect estimate of the harmonic order of a set of partials, the probabilities of which can be computed exactly with the theory.

The virtual pitch theory of Terhardt (1972, 1979) gives an alternative account of the central pitch percept. It is formulated in a deterministic manner, unlike the optimum processor theory, and is schematically illustrated in Figure 8. The theory assumes that spectral frequencies are transformed in the auditory periphery into spectral pitch cues according to cer-

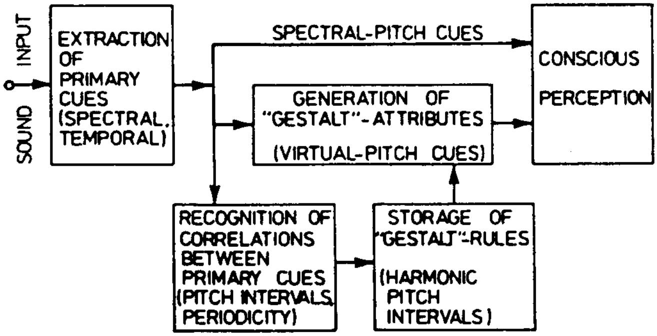

FIGURE 8 Schematic representation of principles underlying the virtual pitch theory of Terhardt (1974b).

tain empirical rules that reflect, for instance, pure-tone pitch shift phenomena discussed in Sections II.B–D. Virtual pitch cues are centrally derived from spectral pitch cues by finding common subharmonics. The model is similar to the optimum processor theory of Goldstein in the sense that both are spectral template matching models. The virtual pitch theory, however, also considers intensities of partials and masking effects. The output of the model is a list of virtual pitch candidates, each with an associated strength, that can be computed from details of the physical stimulus with an algorithm provided by Terhardt (1979).

C. The Role of Unresolved Harmonics

Despite the general development of our understanding that pitch perception is primarily a central process, the question still remains whether totally abandoning Schouten's residue theory is justified. Given the obvious shortcomings of the residue theory it remains true, for instance, that a periodic pulse train retains a certain pitch quality even if all low-order resolvable harmonics have been removed (Moore & Rosen, 1979). Hoekstra (1979) found that the jnd for the missing fundamental of an octave-band wide tone complex remains finite, at about 5 Hz, if the missing fundamental becomes very low and the octave band contains many closely spaced harmonics. Houtsma and Smurzynski (1990) studied pitch identification as well as pitch discrimination performance with complex tones composed of 11 successive harmonics. All complexes had missing fundamental frequencies between 200 and 300 Hz, and harmonic spectra starting between the 7th and the 25th harmonic. Phase relations were either zero (sine) phase, giving a waveform with very distinct peaks, or "negative Schröder" phase (Schröder, 1970) that minimizes the crest factor (peakedness) of the complex-tone signal at the cochlear output. The outcome of the experiments was that, if the number of resolvable harmonics in the complex was progressively reduced, identification performance dropped from near perfect to a low but clearly above-chance level. As shown in Figure 9, jnds increased from about 0.5 to about 5 Hz. The phase relation between harmonics seemed to matter very little. If, on the other hand, the tone complexes contained no resolved harmonics, that is, if the lowest harmonic was on the order 12 or higher, identification and discrimination performance levels remained constant and independent of the harmonic order of the stimulus. Phase, however, turned out to be of great influence on performance, with jnds being almost a factor of 2 larger with the Schröder-phase relation than with the sine-phase relation.

The conclusion from all these experimental results is that not only do low-order harmonics, resolved in the cochlea, contribute to the percept of pitch of a complex tone, but also high-order unresolved harmonics. Their degrees of contribution, however, are quite different. Resolved components

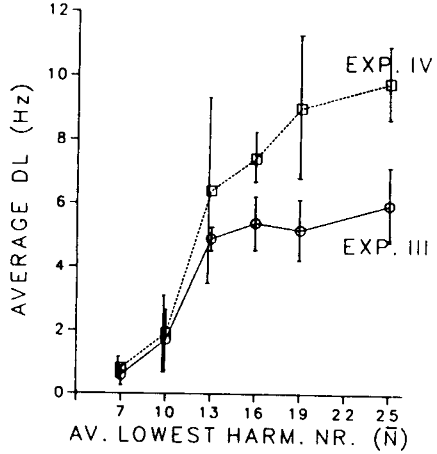

FIGURE 9 Just noticeable differences in fundamental frequency (around 200 Hz) of a complex tone with 11 successive harmonics. The abscissa designates the lowest harmonic number of the 11 harmonics, which are in sine phase (solid curve) or negative "Schröder" phase (dashed curve). Bars indicate standard deviations of mean jnds of four subjects. (From Houtsma and Smurzynski, 1990.)

evoke a stronger, more salient and sharply defined pitch image than unresolved components, as shown by much higher identification scores in, for instance, melodic interval identification tasks and by lower jnds in pitch discrimination tasks. If complex-tone stimuli are broadband, that is, if resolved as well as unresolved partials are present, the latter will generally dominate in determining the perceived pitch, except for complex tones with very low fundamentals (Moore & Peters, 1992).

D. Hybrid Models

In terms of models, one could conclude that two separate neural mechanisms lead to a pitch percept. One operates on neural signals derived from partials, which are resolved in the cochlea, is located centrally because it derives these signals from inputs in the left and right ear simultaneously, and is insensitive to phase relations between complex-tone harmonics. The other operates on temporal properties of cochlear output, similar to the residue mechanism proposed by Schouten (1940). One can also argue, however, that there is only one neural pitch mechanisms in the central auditory system, which yields different performance levels or parameter dependencies for different stimulus conditions. One such model, proposed by Srulovicz and Goldstein (1983), might be a fitting candidate, and it is illustrated in Figure 10. A central spectral magnitude is determined at each frequency by the response of the eighth nerve fiber with characteristic frequency f_c. The

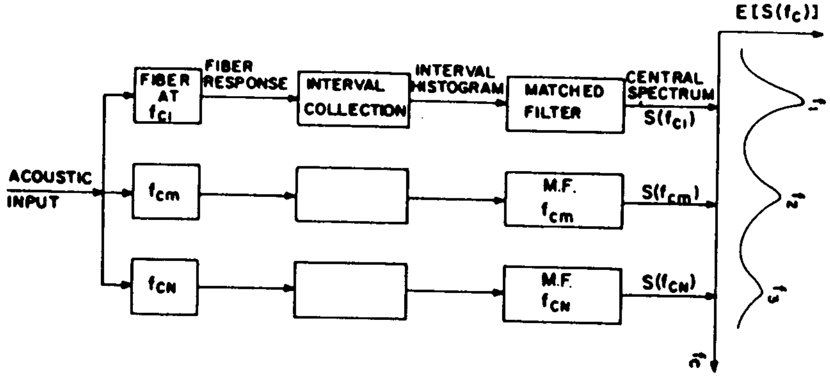

FIGURE 10 Schematic outline of pitch model by Srulovicz and Goldstein (1983).

interspike interval histogram (see Chapter 3) of each fiber is passed through a filter that is matched to its f_c, yielding a single-valued output as contribution to a central spectrum. Frequencies resolved in the cochlea will show up in the central spectrum as well-identified peaks, and a holistic fundamental pitch percept can be derived from this central spectrum in a way described by the optimum processor theory of Goldstein (1973). Degraded pitch performance for unresolved harmonics is predicted by this model because it makes very inefficient use of the abundant fundamental-period information present in the firing patterns of eighth nerve and cochlear nucleus fibers (Horst, Javel, & Farley, 1986; Kim, Rhode, & Greenberg, 1986) for stimuli containing many unresolvable harmonics. Instead of directly computing the inverse of the principal peak in interspike interval histograms, the model maps this phase-locking information at the level of the central spectrum into all the possible harmonics of the fundamental, after which a harmonic-template estimate is made on this central spectrum to find the missing fundamental. This coding and central recovery scheme for the missing fundamental of complex tones with high-order harmonics may seem unnecessarily complicated and very inefficient, but appears consistent with the relatively weak pitch image evoked by such high-order harmonics compared with low-order resolved harmonics.

Another computational model for pitch identification and phase sensitivity with complex-tone stimuli has recently been proposed by Meddis and Hewitt (1991a, 1991b). The model, which is illustrated in Figure 11, has many elements also seen in earlier models (Wightman, 1973; Terhardt, 1972; Srulovicz & Goldstein, 1983), but combines them in a rather unique way. It is composed of (1 and 2) a linear bandpass filter representing the outer and middle ear, (3) a bank of 128 overlapping critical-band (gammatone) filters representing basilar membrane action. It is then followed for

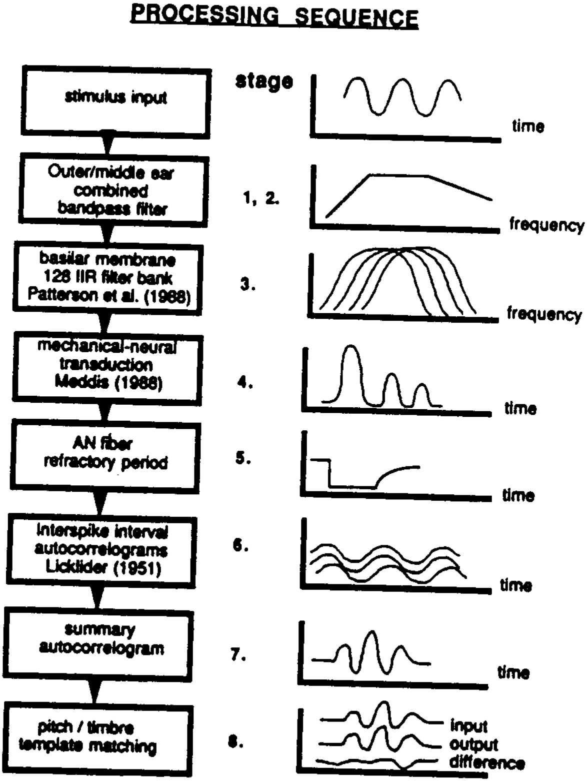

FIGURE 11 Schematic outline of pitch model by Meddis and Hewitt (1991a). Right column shows signal transformations at the various stages. (Reprinted with permission.)

each channel by (4) a hair-cell simulation model (Meddis, 1988), (5) a simple refractory-period model for nerve fibers, and (6) an interspike interval autocorrelation process (Licklider, 1951). Finally, (7) a summary autocorrelogram is formed by averaging the stage 5 output across channels. The pitch percept is represented in this summary autocorrelogram by peak locations, indicating pitch candidates, and peak height indicating relative strength or salience of these pitch candidates.

Although the model incorporates the stochastic nature of neural processes, it uses only their average statistics and is therefore in principle a deterministic model, similar to Wightman's or Terhardt's model. It can identify pitch candidates for any complex stimulus and make a prediction about their relative strengths, at least in an ordinal sense. It is not a discrimination model for making quantitative predictions about pitch confusions, because the output of the model is noiseless and no specific description of a decision model is included. Nevertheless, the model has provided a convincing qualitative account of many known properties of complex-tone pitch, such as the weakening of the sensation with increasing harmonic order, ambiguity, the inharmonic frequency shift effect (Schouten, Ritsma, & Cardozo, 1962), the existence and dominant regions, as well as the effects of repetition pitch and amplitude-modulated noise pitch, which are still to be discussed. To account for dichotic pitch phenomena (Houtsma & Goldstein, 1972; Bilsen & Goldstein, 1974), it seems that the model can be adapted in a fairly simple manner; for instance, by forming the summary autocorrelogram of the last stage from averages across all left and right channels.

E. Pitch of Simultaneous Complex Tones

When listening to music, we are often exposed to complex tones presented simultaneously. If the notes C, E, and G are played simultaneously on three different musical instruments, we can easily perceive the major triad C-E-G. This implies that each of the pitches C, E, and G must be perceived. From an acoustical point of view, the fundamentals of the three notes may be weak or totally absent, and the partials of all three notes are mixed together. Apparently our central auditory system is able to reconstruct groups of harmonically related partials from the total of all resolved partials it receives from the cochlear output.

Beerends and Houtsma (1986, 1989) investigated to what extent our auditory system is able to recognize the two (missing) fundamentals of two simultaneous two-tone harmonic complexes, as a function of the harmonic order of the partials and the manner in which partials were distributed between the ears. They found that deterioration of pitch identification performance with increase in harmonic order was about the same for all presentation conditions; that is, it did not matter very much whether partials of each complex tone went to different ears, all partials went to both ears, or partials of each tone were divided between ears. The conclusion was that frequency information about resolved partials must all end up in the same central pool and that grouping of partials for pitch processing is based on principles other than binaural spatial information (see Chapter 11 for further discussion of this topic). It can be shown that template models like the optimum processor or virtual pitch theories are, in principle, able to account for the observed phenomena (Beerends, 1989).

F. Pitch Ambiguity

Before we leave the topic of tonal pitch, we may wonder whether it is even correct to speak of "the pitch" of a complex tone. On music paper, the sound of a musical instrument is usually represented by a single note, which is thought of as having a certain pitch, duration, and timbre. Laboratory experiments on discrimination, identification, or matching of pitches typically show, however, that the pitch of complex tones can be ambiguous, especially if low-order harmonics are weak or missing or if only a few harmonics are present. Modern pitch theories can adequately account for this ambiguity. Furthermore, another source of ambiguity is, on the one hand, well known but, on the other hand, less well understood and considerably more difficult to model. As in the popular saying about the forest and the trees, the auditory system can perceive a sound complex holistically, where it usually evokes a sensation of a single pitch and some timbre, and also analytically, where it perceives many pitches of individual harmonics or partials. Some models explicitly recognize the existence of holistic and analytic pitch percepts. Terhardt's (1972) virtual pitch theory distinguishes spectral and virtual pitch cues, and Goldstein's (1973) theory distinguishes noisy transformations of resolved frequencies from central estimates of periodicity. None of the theories is able to explain, however, what conditions decide whether analytic or holistic pitch cues are used. Experimental attempts to control and measure conditions for analytic or holistic pitch perception (Smoorenburg, 1970; Houtsma & Fleuren, 1991) generally show that it is difficult to control the perception mode by experimental conditions in individual listeners. Some listeners have a strong inclination toward analytic perception behavior, others show a strong tendency toward holistic behavior, and still others show inconsistent behavior. Only group-averaged behavior—for instance, under a condition where holistic and analytic perception modes lead to opposite responses—shows some definite tendencies. The most important one is that, for complex tones with high-order harmonics, listeners' responses tend to divide about half-and-half into analytic and synthetic responses, whereas for tones with low-order harmonics analytic responses dominate. Lowering the harmonic order of a complex tone enhances both the holistic pitch percept, because of the dominant region effect (Ritsma, 1967), and the analytic pitch percept, because of the increased spectral resolution in the auditory periphery (see Chapter 5). Apparently, the effect of the latter is stronger than that of the former, at least with two-tone complexes. Much more systematic experimental evidence is required, however, before serious modeling attempts can be undertaken to describe the precise relationship between analytic and holistic pitch perception behavior.

IV. NONTONAL PITCH

A. Repetition Pitch

While visiting the French castle of Chantilly de la Cour, the Dutch physicist Christiaan Huygens noticed that the garden fountain, located in a vertical recess surrounded by marble steps, produced a noisy sound with a distinct musical pitch. Huygens described the pitch as corresponding with the sound of an open organ pipe of a length matching the depth of the stairs. Since that observation in 1693 it has been found that, in general, if an arbitrary sound $s(t)$ and its echo $s(t - T)$ are added together, a repetition pitch is heard that corresponds with a pure tone of frequency $1/T$. The sound $s(t)$ may be a simple click, a burst of white noise, a sample of speech, or just about any other broadband sound. The effect has been studied systematically for monotic and diotic conditions, where $s(t)$ and $s(t - T)$ go either to one ear or to both ears (Bilsen, 1968; Yost, Hill, & Perez-Falcon, 1978; Yost & Hill, 1978), and also for dichotic conditions, where the signals $s(t)$ and $s(t - T)$ go to different ears (Bilsen & Goldstein, 1974). Repetition pitch effects are typically found for delay times between 1 and 10 ms, yielding pitches varying from 100 to 1000 Hz. The effect is even stronger if there are many repeated echoes, for instance a signal $s(t) = x(t) + a_1x(t - T) + a_2x(t - 2T) + \ldots + a_nx(t - nT)$. Such signals are often referred to as *comb-filtered signals,* because the repeated echoes in the time domain cause more sharply defined maxima to occur in the spectrum at frequencies $f_n = n/T$ ($n = 0, 1, 2, 3, \ldots$). For this reason repetition pitch phenomena can, at least in principle, be accounted for by the same models used to describe the pitch of complex tones. Other models that are based on the interaural cross-correlation between similarly tuned channels have been developed by Blauert (1974) and Bilsen (1977).

The repetition pitch phenomenon can sometimes be used very creatively for special effects in electronic music. It can also be a nuisance. For instance, in a concert hall, a wrongly placed wall or other reflecting surface may cause at a particular seat an echo with a delay between 1 and 10 ms, producing an audible sound coloration.

B. Huggins Pitch and Edge Pitch

There are other conditions under which broadband noise can evoke a pitch sensation. One of these conditions is known in the literature as *Huggins pitch* (Cramer & Huggins, 1958). It arises if broadband noise signals are dichotically presented, identical in every respect except for an interaural phase shift over a small frequency region below 1500 Hz. A faint pitch is heard that appears to correspond to the center frequency of the phase shift region. The phenomenon is regarded as evidence of an interaural subtraction process such as, for instance, in the equalization and cancellation model of Durlach (1972) discussed in Chapter 10. The faint pitch is then attributed to a central narrow band of noise, which remains after the subtraction process.

If broadband noise is filtered, either highpass or lowpass, and if the filter's cutoff slope is sufficiently steep, a vague pitch is heard at or around the spectral edge (Small & Daniloff, 1967; Fastl, 1971). If the spectrum of a complex tone is abruptly terminated at some harmonic of high order, a much more pronounced pitch is evoked near the spectral edge frequency, which can even be comparable in salience to that of a pure tone if the phase spectrum is optimal (Kohlrausch & Houtsma, 1992). Although both pitch phenomena deal with sensations associated with some spectral discontinuity, they behave quantitatively in such different ways that they are probably based on entirely different auditory mechanisms (Kohlrausch & Houtsma, 1992).

The noise edge pitch described by Small and Daniloff and by Fastl can also be created dichotically. Klein and Hartmann (1981) described how, if both ears are stimulated with the same broadband noise signal, except for an interaural phase transition function that steps from 0 to 180° at a frequency f_c, two faint pitches are heard, one slightly below and the other slightly above the phase transition frequency f_c. Frijns, Raatgever, and Bilsen (1986) found a fairly unimodal distribution around f_c for this faint pitch. Binaural edge pitch is not very salient. Subjects can tell to some extent whether one sensation is higher or lower than another or match the binaural stimulus in pitch to a pure tone with some degree of accuracy. It has never been shown, however, that melodies or melodic pitch intervals can be recognized if the phase step frequency is given discrete values on a musical scale.

C. Pitch of Amplitude-Modulated Noise

If broadband noise is periodically gated or amplitude modulated by a sine wave, a pitchlike phenomenon associated with the envelope periodicity is observed. Miller and Taylor (1948) showed that subjects could discriminate noise-gating frequencies below 100 Hz with about the same precision as they could pure tones. This is shown in Figure 12. Beyond 100 Hz, the jnds for the interruption rate become much larger than pure-tone jnds, the whole interruption-pitch phenomenon fading away at about 300 Hz. Similar to binaural edge pitch, it is not entirely clear whether AM or interruption pitch is a true pitch phenomenon, especially in the musical sense. The fact that subjects can discriminate between high and low pitches, as in Miller and Taylor's experiment, suggests that the percept satisfies the official definition of pitch. Moreover, Burns and Viemeister (1976) have shown that subjects could identify seven known melodies played with AM noise in the modulation-frequency range of 100–200 Hz at a 90% correct level. The same authors found, however, just as Houtsma, Wicke, and Ordubadi (1980) did, that musically trained subjects were not able to score better than about 50% correct if asked to identify random melodic intervals that differed by semitone steps.

The fact that periodically modulated noise appears to evoke pitch sensations has traditionally been regarded as direct evidence that pitch is medi-

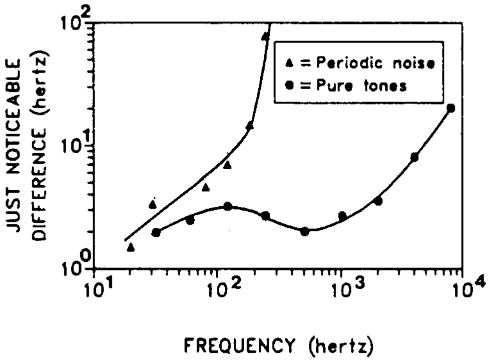

FIGURE 12 Just noticeable differences in the interruption rate of white noise (circles) and the frequency of a pure tone (triangles). (After Miller and Taylor, 1948.)

ated by temporal mechanisms in the auditory system, because the long-term average spectra of these signals are flat. Pierce, Lipes, and Cheetham (1975) have shown, however, that this argument does not necessarily hold, because short-term spectra do contain information about the modulation or interruption frequency. Houtsma et al. (1980) compared predictions by both temporal and short-term spectral models with measured pitch recognition scores and found that measured behavior generally supports the temporal view.

V. PITCH SCALES: RELATIVE AND ABSOLUTE

A. Relative Pitch

The role of pitch in music is based primarily on pitch relations and not on absolute pitch values. Sets of notes that people have used to make music throughout history, from the Greek tetrachords and Gregorian church modes to present-day major, minor, and chromatic scales, all have well-defined mutual relationships without the necessity for an absolute reference. When musicians play in an ensemble or sing together, it is usually sufficient that they tune their scales to one another, with the instruments that are most difficult to tune (such as piano or organ) being taken as the reference. The international convention to fix the fundamental frequency of the A_4 (the A in the fourth octave on a piano keyboard) at 440 Hz is only of rather recent origin, is typical only of Western music, and is still not endorsed by some of our major symphony orchestras.

Given the rather low priority obviously placed on absolute frequency standards in music, it will not be surprising that pitch perception is relation oriented rather than absolute. Many people, for instance, are able to sing a tune without knowing the key in which they are singing. Melodic steps, that is, sequential frequency ratios, can be produced with great accuracy without the necessity or even the awareness of any absolute reference. The first formal musical training children receive in school is the do-re-mi scale, which is relative: all tone steps relative to the note "do" are well defined, but "do" itself can be taken as any convenient frequency.

Formal musical pitch perception training, called *solfeggio,* consists mainly of strengthening and formalizing a natural ability to recognize, memorize, and reproduce certain sets of frequency ratio steps. Music students learn to associate names such as octave, fifth, or minor third with simultaneous or sequential sounds they hear, and they learn to sing melodic intervals from written music notation. In this way an *absolute* sense of *relative* pitch is developed, which is considered to be a standard skill of every professional musician.

Musical scales are in principle built from arbitrary frequency steps. Our modern diatonic and chromatic scales represent only a particular historical development in our Western culture. Scales used in other cultures often contain intervals quite different from those in Western tone scales. Some basic intervals, however, for instance the octave (ratio 2:1) and fifth (ratio 3:2), are found in many of the non-Western scales, probably because they occur between clearly audible elements (the first, second, and third harmonics) in natural periodic sounds. Especially if music is polyphonic or harmonic, which has been a characteristic of Western music since the twelfth century, the necessity of avoiding beats between partials forces one to choose melodic scale steps that are matched as well as possible to the frequency ratios of partials in instrumental or vocal sounds. The fact that a perfect match is mathematically impossible and therefore compromises must be made has led to the development of various tuning systems or *temperaments,* such as the Pythagorean, just or natural, mean-tone, and equal temperament. The introduction of alternative tone systems in this century, for instance, the quarter-tone system (Hába, 1927) or the 31-tone system (Fokker, 1949), have all met with limited success because of the unpleasant-sounding beats that occur between mistuned partials. Such beats do not occur with the tuning system proposed by Mathews and Pierce (1980), where the frequencies of partials of the sounds to be used are chosen to match the novel melodic tone steps obtained, for instance, by dividing an octave into 13 equal-ratio frequency steps. Such sounds may be difficult to find in the natural world, but can easily be synthesized with modern digital techniques.

A developed absolute sense of relative pitch can also be used for the psychoacoustical study of pitch perception. Requiring trained subjects to identify or reproduce aurally presented musical intervals or short tone sequences with experimental test sounds is a good alternative to the more conventional techniques of pitch matching or low–high discrimination, especially if it is not clear from the start that the sensation being studied is a real pitch phenomenon (Houtsma, 1984).

B. Pitch Contours in Speech

When we speak, our voice produces either periodic sounds for vowels and voiced consonants or noisy and aperiodic sounds for fricatives and unvoiced stop consonants. With voiced sounds our vocal cords vibrate at a rate f_0, which is therefore the fundamental of the vowel or voiced-consonant sound. During a spoken sentence the value of f_o varies with time, forming a more or less continuous pattern. This so-called *intonation pattern* carries important prosodic information and follows very specific language-dependent rules (Hart, Collier, & Cohen, 1990).

One might wonder whether, from a perceptual viewpoint, pitch interval relationships are the same in running speech as they are in music. One might expect this to be the case because a spoken vowel is, in principle, no different from any other musical sound and the human voice is, after all, the most frequently used musical instrument. On the other hand, there is a clear categorical difference between f_0 contours in speech and in music. The former are always continuous, the latter almost always discrete and restricted to a limited number of f_0 values on the chosen musical scale.

Hermes and van Gestel (1991) have found evidence that pitch relations in speech and in music are perceptually different. They presented subjects with two "ma-*ma*-ma" utterances, each in a different octave range, with the middle syllable being accented by making an up–down f_0 sweep. One pitch accent was fixed by the experimenter, while the other could be adjusted by subjects to match the prominence of the accent. Analysis of the matched frequency excursions showed that equal accent prominence was not given by equal frequency or log-frequency excursions, but rather by equal excursions along an equivalent rectangular band (ERB) scale as discussed in Chapter 5. This implies that the prominence of a pitch accent in speech is determined by the number of critical bands the fundamental f_0 is swept through, a thought that was also the basis of the mel scale of Stevens and Zwicker discussed in Section II.A. An unresolved problem with this notion, however, is that the mel scale was intended for pure tones, whereas the pitch of most speech signals is mostly virtual, with very little energy in the fundamental frequency component.

C. Absolute Pitch

Absolute pitch refers to the ability of some people to identify musical sounds by their proper note name, or to name the key of a piece of music, without the use of any obvious external reference. Despite the rather large number of studies that have been devoted to this topic, information is still mostly empirical and sketchy, and our understanding of the phenomenon is still rather poor.

Perhaps the most systematic and comprehensive studies on the topic were done by Bachem (1937, 1940, 1954). Among people appearing to possess absolute pitch he distinguishes between genuine and acquired absolute pitch skills. Possessors of *genuine* absolute pitch typically make quick absolute identifications, accurate within a semitone, with octave confusions being the principal source of errors. *Acquired* skills are behaviorally characterized by slow judgments, as if subjects are trying to recall some learned reference like the A_4 for orchestra musicians or an extreme of the vocal range for singers. Given enough time, these subjects can make fairly accurate absolute pitch judgments, but if forced to respond quickly they will typically make large errors.

Bachem (1954) measured free-field pure-tone frequency jnds as a function of the temporal separation between tones which varied between one second and one week. Figure 13 shows jnds expressed as a percentage of the

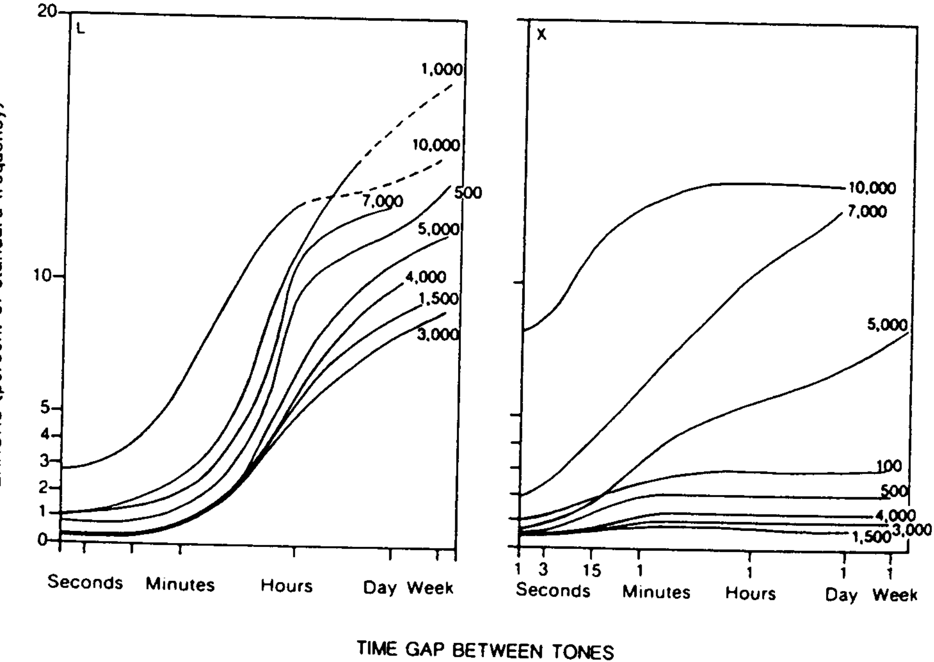

FIGURE 13 Just noticeable differences for pure-tone frequency as a function of the inter-tone time gap from two subjects. Duration of tones was 2 s and presentation was freefield. Subject *X* claimed absolute pitch, subject *L* did not. (Data from Bachem, 1954, reprinted with permission.)

tone's frequency of two subjects, one without (*L*) and the other with absolute pitch (*X*). Subject *L* shows a jnd that grows steadily with time, indicating a degrading memory trace for the pitch of the first note. Subject *X* shows a fairly constant jnd of about 3%, at least for frequencies below 5000 Hz, independent of how much time has lapsed between tones. This subject apparently labels the perceived pitches with some verbal code and ultimately compares the labels. Labeling is done by this subject with an accuracy of 3%.

It is still not known whether the observed behavioral differences between those who do and do not possess absolute pitch reflect actual physiological differences. There is some indication (Bachem, 1940) that absolute pitch requires an innate ability, combined with the right exposure during a critical development period at an early age.

VI. MULTIDIMENSIONAL ASPECTS OF PITCH

Up to this point the attribute pitch has been treated in this chapter as a one-dimensional entity. This seems to be in accordance with the ANSI (1973) definition, which describes pitch as a sensory attribute that enables ordering of sounds on a scale extending from low to high. It also appears to be consistent with the centuries-old practice of staff notation, where pitches of musical sounds are represented by the places of musical notes on a staff. Such a staff is actually nothing other than a visually convenient graphical representation of objects on a one-dimensional scale.

If one looks at musical practice and sees how pitch and pitch relationships are treated in music theory and composition analysis, however, one can hardly avoid drawing the conclusion that there must be more than a single dimension to the sensation of pitch. If, for instance, one takes the dimension that underlies conventional music notation as the only dimension of pitch, one has difficulty explaining why a C_5 sounds closer to a C_4 than an $F^{\#}_4$. If one runs up or down a diatonic or chromatic scale, one clearly perceives a circularity, where in every octave pitches seem to repeat themselves in some sense. One thereafter might want a perceptual representation of pitch in which stimuli that are close together sound similar, and stimuli that are far apart sound dissimilar.

Octave similarity has formed the basis of the two-dimensional helix representation of pitch shown in Figure 14(a) (Revesz, 1954; Shepard, 1964). Going around one turn of the helix, one traverses the 12 chromatic pitches in an octave (C, C#, D, etc.). Completion of one turn brings one almost but not quite back to the place of origin. The dimension one varies by going around the helix is often called *pitch chroma*, and the axial distance one travels is referred to as *tone height*. The degree of octave similarity is represented by the amount of stretch in the helix along its axis. In the extreme case of a

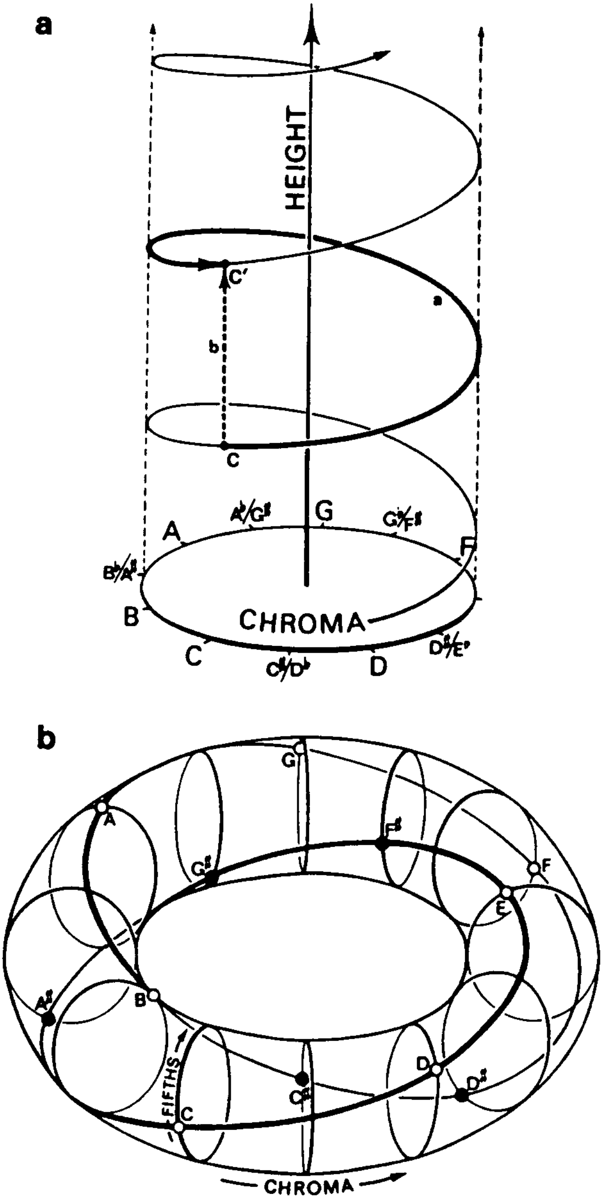

FIGURE 14 Multidimensional representations of pitch. (a) Simple helix representing pitch chroma and tone height. (b) Torus with double helix, representing the chroma circle and the circle of fifths. (From Shepard, 1982, reprinted with permission.)

completely stretched helix, we have no octave similarity and have, in fact, returned to the one-dimensional representation which was discussed earlier. The other extreme of the helix being compressed to a circle yields a pitch scale with octave identity. Shepard (1964) has shown how to make stimuli that have this sensory circular property. Stimuli moving around the circle evoke a sensation of an infinitely rising or falling pitch. These and similar stimuli have been used by composers as special sound effects and have occasionally been used as research tools in psychoacoustic experiments (Shepard, 1964; Allik, Dzhafarov, Houtsma, Ross, & Versfeld, 1989; Deutsch, 1991).

In addition to chromatic distance and octave similarity are other principles in music theory that determine proximity or distance between pitches. The circle of fifths, for instance, recognizes the close relationship between notes that are a fifth apart and have a tonic-dominant or tonic-subdominant relationship. Harmony in traditional Western music is based on this principle, and one can find it directly, for instance, in the physical layout of the bass keys of an accordion. If one combines the principles of the chroma circle and the circle of fifths, one obtains the pitch representation of a double helix wrapped around a torus, shown in Figure 14(b). These and other even more complex pitch representations can be found in a review chapter on this topic by Shepard (1982).

Many of the principles other than chromatic distance, particularly octave similarity and the circle of fifths, are of harmonic rather than melodic origin. They follow from frequency relationships that are found between overtones of natural periodic sounds such as the human voice, strings, or wind instruments. The relevance of such principles should therefore depend heavily on the spectral composition of sounds used to evoke pitch sensations. A multidimensional pitch space for harmonic sounds may be very different from a pitch space for sounds like church bells (Houtsma & Tholen, 1987) or sounds with stretched partials (Mathews & Pierce, 1980). There is therefore good reason to have doubts about the general validity of any abstract theory of multidimensional pitch space that does not deal with the issue of orchestration.

References

Allik, J., Dzhafarov, E. N., Houtsma, A. J. M., Ross, J., & Versfeld, N. J. (1989). Pitch motion with random chord sequences. *Perception and Psychophysics, 46,* 513–527.

American National Standards Institute. (1973). American national psychoacoustical terminology. S3.20. New York: Author.

Bachem, A. (1937). Various types of absolute pitch. *Journal of the Acoustical Society of America, 9,* 146–151.

Bachem, A. (1940). The genesis of absolute pitch. *Journal of the Acoustical Society of America, 11,* 434–439.

Bachem, A. (1954). Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America, 26,* 751–753.

Beerends, J. G. (1989). Pitches of simultaneous complex tones. Ph.D. dissertation, Eindhoven University of Technology.

Beerends, J. G., & Houtsma, A. J. M. (1986). Pitch identification of simultaneous dichotic two-tone complexes. *Journal of the Acoustical Society of America, 80,* 1048–1056.

Beerends, J. G., & Houtsma, A. J. M. (1989). Pitch identification of simultaneous diotic and dichotic two-tone complexes. *Journal of the Acoustical Society of America, 85,* 813–819.

Bilsen, F. A. (1968). On the interaction of a sound with its repetitions. Ph.D. thesis, Delft, Waltman.

Bilsen, F. A. (1977). Pitch of noise signals: Evidence for a "central spectrum." *Journal of the Acoustical Society of America, 61,* 150–161.

Bilsen, F. A., & Goldstein, J. L. (1974). Pitch of dichotically delayed noise and its possible spectral basis. *Journal of the Acoustical Society of America, 55,* 292–296.

Blauert, J. (1974). *Räumliches Hören.* Stuttgart: Hirzel Verlag.

Burns, E. M., & Viemeister, N. F. (1976). Nonspectral pitch. *Journal of the Acoustical Society of America, 60,* 863–869.

Brink, G. van den (1970). Experiments on binaural diplacusis. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing* (pp. 362–374). Leiden: Sijthoff.

Cramer, E. M., & Huggins, W. H. (1958). Creation of pitch through binaural interaction. *Journal of the Acoustical Society of America, 30,* 413–417.

Deutsch, D. (1991). Pitch proximity in the grouping of simultaneous tones. *Music Perception, 9,* 185–198.

Durlach, N. I. (1972). Binaural signal detection: equalization and cancellation theory. In J. V. Tobias (Ed.), *Foundations of modern auditory theory* (Vol. 2, pp. 369–462). New York: Academic Press.

Fastl, H. (1971). Über Tonhöhenempfindungen bei Rauschen. *Acustica, 25,* 350–354.

Fletcher, H. (1924). The physical criterion for determining the pitch of musical tone. *Physical Review, 23,* 427–437.

Fokker, A. D. (1949). *Just intonation.* The Hague: M. Nijhoff.

Frijns, J. H. M., Raatgever, J., & Bilsen, F. A. (1986). A central spectrum theory of binaural processing. The binaural edge pitch revisited. *Journal of the Acoustical Society of America, 80,* 442–451.

Goldstein, J. L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America, 54,* 1496–1516.

Hába, A. (1927). *Neue Harmonielehre des diatonischen, chromatischen, Viertel-, Drittel-, Sechstel- und Zwölfteltonsystems.* Leipzig: Kirstner and Siegel.

Hart, J.'t, Collier, R., & Cohen, A. (1990). *A perceptual study of intonation.* Cambridge: Cambridge University Press.

Hartmann, W. M. (1978). The effect of amplitude envelope on the pitch of sinewave tones. *Journal of the Acoustical Society of America, 63,* 1105–1113.

Helmholtz, H. L. F. von (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik,* Braunschweig: F. Vieweg & Sohn. (English trans. A. J. Ellis; New York: Dover, 1954).

Hermes, D. J., & van Gestel, J. C. (1991). The frequency scale of speech intonation. *Journal of the Acoustical Society of America, 90,* 97–102.

Hoekstra, A. (1979). Frequency discrimination and frequency analysis in hearing. Ph.D. thesis, University of Groningen.

Horst, J. W., Javel, E., & Farley, G. R. (1986). Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms. *Journal of the Acoustical Society of America, 79,* 398–416.

Houtsma, A. J. M. (1984). Pitch salience of various complex sounds. *Music Perception, 1,* 296–307.

Houtsma, A. J. M., & Fleuren, J. F. M. (1991). Analytic and synthetic pitch of two-tone complexes. *Journal of the Acoustical Society of America, 90,* 1674–1676.

Houtsma, A. J. M., & Goldstein, J. L. (1972). The central origin of the pitch of complex tones: Evidence from musical interval recognition. *Journal of the Acoustical Society of America, 51,* 520–529.

Houtsma, A. J. M., & Smurzynski, J. (1990). Pitch identification and discrimination for complex tones with many harmonics. *Journal of the Acoustical Society of America, 87,* 304–310.

Houtsma, A. J. M., & Tholen, H. J. G. M. (1987). A carillon of major-third bells: II. A perceptual evaluation. *Music Perception, 4,* 255–266.

Houtsma, A. J. M., Wicke, R. W., & Ordubadi, A. (1980). Pitch of amplitude-modulated low-pass noise and predictions by temporal and spectral theories. *Journal of the Acoustical Society of America, 67,* 1312–1322.

Kim, D. O., Rhode, W. S., & Greenberg, S. R. (1986). Responses of cochlear nucleus neurons to speech signals: Neural encoding of pitch, intensity and other parameters. In B. C. J. Moore & R. D. Patterson (Eds.), *Auditory frequency selectivity* (pp. 281–288). London: Plenum Press.

Klein, M. A., & Hartmann, W. M. (1981). Binaural edge pitch. *Journal of the Acoustical Society of America, 70,* 51–61.

Kohlrausch, A., & Houtsma, A. J. M. (1992). Pitch related to spectral edges of broadband signals. *Philosophical Transactions of the Royal Society of London, 336* Part B, 375–382.

Licklider, J. C. R. (1951). A duplex theory of pitch perception. *Experientia, 7,* 128–133.

Mathews, M. V., & Pierce, J. R. (1980). Harmony and nonharmonic partials. *Journal of the Acoustical Society of America, 68,* 1252–1257.

Meddis, R. (1988). Simulation of auditory-neural transduction: Further studies. *Journal of the Acoustical Society of America, 83,* 1056–1063.

Meddis, R., & Hewitt, M. J. (1991a). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification. *Journal of the Acoustical Society of America, 89,* 2866–2882.

Meddis, R., & Hewitt, M. J. (1991b). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II. Phase sensitivity. *Journal of the Acoustical Society of America, 89,* 2883–2894.

Miller, G. A., & Taylor, W. G. (1948). The perception of repeated bursts of noise. *Journal of the Acoustical Society of America, 20,* 171–182.

Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *Journal of the Acoustical Society of America, 54,* 610–619.

Moore, B. C. J., & Peters, R. W. (1992). Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity. *Journal of the Acoustical Society of America, 91,* 2881–2893.

Moore, B. C. J., & Rosen, S. M. (1979). Tune recognition with reduced pitch and interval information. *Quarterly Journal of Experimental Psychology, 31,* 229–240.

Morgan, C. T., Garner, W. R., & Galambos, R. (1951). Pitch and intensity. *Journal of the Acoustical Society of America, 23,* 658–663.

Ohm, G. W. (1843). Über die Definition des Tones nebst daran geknupfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen. *Annalen für Physik und Chemie, 59,* 513–565.

Pierce, J. R., Lipes, R., & Cheetham, C. (1975). Uncertainty concerning the direct use of time information in hearing: Place clues in white-spectra stimuli. *Journal of the Acoustical Society of America, 61,* 1609–1621.

Plomp, R. (1967). Pitch of complex tones. *Journal of the Acoustical Society of America, 41,* 1526–1533.

Revesz, G. (1954). *Introduction to the psychology of music.* Norman: University of Oklahoma Press.

Ritsma, R. J. (1962). The existence region of the tonal residue I. *Journal of the Acoustical Society of America, 34,* 1224–1229.

Ritsma, R. J. (1967). Frequencies dominant in the perception of the pitch of complex sounds. *Journal of the Acoustical Society of America, 42,* 191–198.

Rossing, T. D. and Houtsma, A. J. M. (1986). Effects of signal envelope on the pitch of short sinusoidal tones. *Journal of the Acoustical Society of America, 79,* 1926–1933.

Schouten, J. F. (1938). The perception of subjective tones. *Proceedings of the Koninklijke Akademie van Wetenschap, 41,* 1086–1093.

Schouten, J. F. (1940). The residue and the mechanism of hearing. *Proceedings of the Koninklijke Akademie van Wetenschap, 43,* 991–999.

Schouten, J. F., Ritsma, R. J., & Cardozo, B. L. (1962). Pitch of the residue. *Journal of the Acoustical Society of America, 34,* 1418–1424.

Schröder, M. R. (1970). Synthesis of low-peak-factor signals and binary sequences with low autocorrelation. *IEEE Transactions on Information Theory, IT-16,* 85–89.

Seebeck, A. (1841). Beobachtungen über einige Bedingungen der Entstehung von Tönen. *Annalen für Physik und Chemie, 53,* 417–436.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36,* 2346–2353.

Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 343–390). New York: Academic Press.

Shower, E. G., & Biddulph, R. (1931). Differential pitch sensitivity of the ear. *Journal of the Acoustical Society of America, 3,* 275–287.

Siebert, W. M. (1970). Frequency discrimination in the auditory system: Place or periodicity mechanisms. *Proceedings of the IEEE, 58,* 723–730.

Small, A. M., Jr., & Daniloff, R. G. (1967). Pitch of noise bands. *Journal of the Acoustical Society of America, 41,* 506–512.

Smoorenburg, G. F. (1970). Pitch perception of two-frequency stimuli. *Journal of the Acoustical Society of America, 48,* 924–941.

Srulovicz, P., & Goldstein, J. L. (1983). A central spectrum model: A synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. *Journal of the Acoustical Society of America, 73,* 1266–1276.

Stevens, S. S. (1935). The relation of pitch to intensity. *Journal of the Acoustical Society of America, 6,* 150–154.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude of pitch. *Journal of the Acoustical Society of America, 8,* 185–190.

Terhardt, E. (1972). Zur Tonhöhewahrnehmung von Klängen II: Ein Funktionsschema. *Acustica, 26,* 187–199.

Terhardt, E. (1974a). Pitch of pure tones: Its relation to intensity. In E. Zwicker and E. Terhardt (Eds.), *Facts and models in hearing* (pp. 353–360). Stuttgart: Springer Verlag.

Terhardt, E. (1974b). Pitch, consonance and harmony. *Journal of the Acoustical Society of America, 55,* 1061–1069.

Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research, 1,* 155–182.

Terhardt, E., & Fastl, H. (1971). Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen. *Acustica, 25,* 53–61.

Verschuure, J., & van Meeteren, A. A. (1975). The effect of intensity on pitch. *Acustica, 32,* 33–44.

Wier, C. C., Jesteadt, W., & Green, D. M. (1977). Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America, 61,* 178–184.

Wightman, F. L. (1973). The pattern-transformation model of pitch. *Journal of the Acoustical Society of America, 54,* 407–416.

Yost, W. A., & Hill, R. (1978). Strength of the pitches associated with ripple noise. *Journal of the Acoustical Society of America, 64,* 485–492.

Yost, W. A., Hill, R., & Perez-Falcon, T. (1978). Pitch and pitch discrimination of broadband signals with rippled power spectra. *Journal of the Acoustical Society of America, 63,* 1166–1173.

Zwicker, E., & Feldtkeller, R. (1967). *Das Ohr als Nachrichtenempfänger.* Stuttgart: S. Hirzel Verlag.