Chapter 13
Pitch

Pitch is the psychological sensation of the highness or the lowness of a tone. Pitch is the basis of melody in music and of emotion in speech. Without pitch, music would consist only of rhythm and loudness. Without pitch, speech would be monotonic—robotic. As human beings, we have astonishingly keen perception of pitch. The principal physical correlate of the psychological sensation of pitch is the physical property of frequency, and our keen perception of pitch allows us to make fine discriminations along a frequency scale. Between 100 and 10,000 Hz we can discriminate more than 2,000 different frequencies!

13.1 Pitch of Sine Tones: Place Theory

Because frequency is so important to pitch, we begin the study of pitch with a sine tone, which has only a single frequency. From Chap. 11 you know that the basilar membrane is selective for frequency. Different frequencies cause the membrane to vibrate maximally at different locations. This, in turn, causes particular hair cells to be maximally excited by particular frequencies. Each hair cell is connected to particular neurons in the auditory nerve, and this establishes a neural selectivity for frequency that is maintained throughout the entire auditory system. The frequency selectivity of neurons is called "tonotopic analysis," *tono-* for tone frequency, and *topic* for location, meaning that different frequencies excite slightly different locations in the brain.

Nothing could be more natural than to assume that the sensation of pitch is the result of tonotopic analysis. This idea is called the "place theory" of pitch perception. According to this theory, you hear different pitches because different neurons (different places in your brain) are excited. From years of research, including recent studies of persons with cochlear implants, we know that this place theory is capable of accounting for many aspects of pitch perception. However, there is more…

W.M. Hartmann, *Principles of Musical Acoustics*, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-1-4614-6786-1_13,

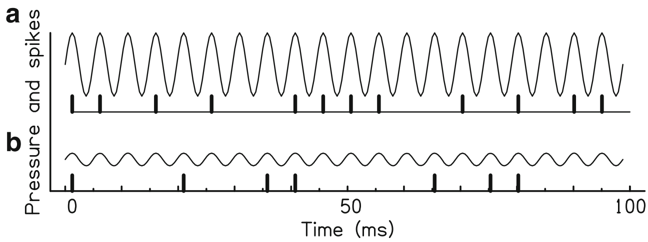

Fig. 13.1 The intense tone in (**a**) produces more neural spikes than the weak tone in (**b**), but both tones produce synchronized spikes

13.2 Pitch: Timing Theory

You know about neural firing rate. You know that a greater firing rate corresponds to greater loudness. You know that a high firing rate in certain neurons encodes pitch—the place theory mechanism. Now you need to recall a third property of the neural code, introduced in Chap. 11, namely *neural synchrony.* The firing of a hair cell and the spikes in the neurons connected to it synchronize to the input tone. This happens because a hair cell tends to fire on a particular phase of the pressure wave in the cochlea. Figure 13.1 shows the firing of a hair cell that fires on waveform peaks. Part **a** shows 1/10 s of a 200-Hz sine tone of moderate intensity. This tone causes 12 neural spikes, so the firing rate is about 120 spikes per second, but the spikes don't come at random times. The time intervals between spikes are integer multiples of a period, specifically: 5, 10, 10, 15, 5, 5, 5, 15, 10, 10, 5 ms. This sequence is only an example. In the next 1/10 s, the sequence of intervals will no doubt be different, but all the intervals will be multiples of 5 ms.

In this way, the timing of the neural spikes encodes the period of the tone, and this provides another cue to the brain for the pitch of the tone. The idea that the synchrony of neural spikes leads to the sensation of pitch is called the "timing theory" of pitch perception. Part **b** of Fig. 13.1 shows another 200-Hz tone, but less intense. There are only seven spikes in 1/10 s, but they too are synchronized with the signal.

So far as the timing theory is concerned, the intervals are still multiples of 5 ms and the timing information about pitch is not different for this weaker tone.

13.3 Pitch of a Complex Tone

Most sounds that we encounter are complex. They have multiple frequency components. If the tone is periodic then the components are harmonic—their frequencies are integer multiples of a fundamental frequency. Figure 13.2 shows a very typical tone. It could be a musical instrument tone or a sustained spoken vowel. Ten

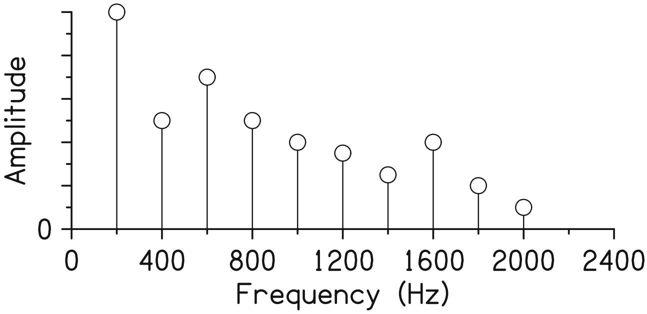

Fig. 13.2 A complex tone with ten harmonics of 200 Hz

harmonics are shown. There are more, but they are normally weaker than the first half dozen. The fundamental is the strongest harmonic, which is typical, but for many other complex tones the fundamental is not the strongest. In fact, the fundamental might not even exist. Nothing that we say about pitch requires that the fundamental be strong or even that it be present.

The tone in Fig. 13.2 has a pitch of 200 Hz. That means that if you listen to that tone for about a second or less, and try to sing what you heard, you will sing a tone with a fundamental frequency of 200 Hz. Or, to be more precise, you might have a sine-tone generator with a knob that controls the frequency. When you adjust the sine frequency to match the pitch of the complex tone it turns out that you set the generator to 200 Hz.

This experimental result makes the timing theory look very good. Because the components are all harmonics of 200 Hz, the period of the waveform is 1/200 s. The prediction of the timing theory and the experimental results agree that (1) the pitch depends on the period and does not depend much on the relative strengths of the harmonics. (2) The pitch does not depend on the fundamental or other low harmonics being present in the spectrum. For instance, a tone with the spectrum of Fig. 13.3 also has a period of 1/200 s and it also leads to a pitch of 200 Hz. (3) The pitch is insensitive to the overall intensity of the tone. By contrast, it's not clear what the place theory has to say about a complex tone. The place theory might predict that a tone with ten harmonics, like Fig. 13.2, should lead to ten different pitches. After all, the ten different frequencies in the tone excite the basilar membrane at ten different places.

Before giving up on the place theory though, we need to do some more experimenting. Suppose we have a spectrum like Fig. 13.4 with a 200-Hz fundamental component and a 10th harmonic component at 2,000 Hz and no other components. Once again, the period is 1/200 s, but this time one clearly hears two distinct pitches, one at 200 Hz and the other at 2,000. This observation agrees with place theory and not with timing theory. The obvious conclusion is that neither the place theory of pitch perception nor the timing theory of pitch perception is good enough to account for the sensations produced by all possible tones. We shall continue to try to find a better theory.

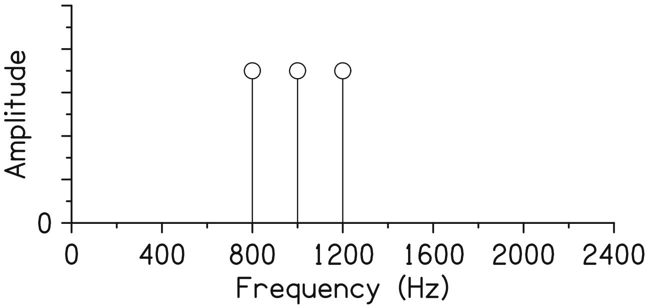

Fig. 13.3 The spectrum has components at 800, 1,000, and 1,200 Hz, corresponding to the 4th, 5th, and 6th harmonics of 200 Hz. According to most listeners, the pitch is 200 Hz

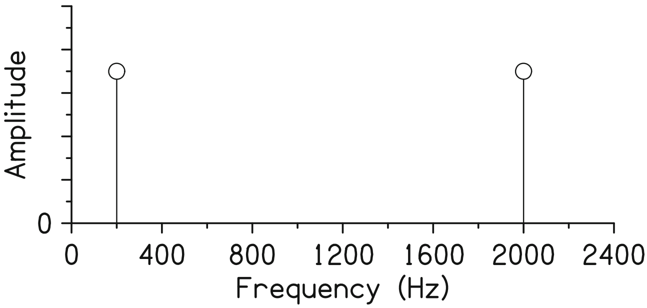

Fig. 13.4 The spectrum has components at 200 and 2,000 Hz, corresponding to the first and 10th harmonics of 200 Hz. The two components are heard separately—analytic listening

13.4 The Template Theory

The template theory (or template model) of pitch perception deals with some of the problems of the place theory. The template theory says that from early experience (maybe even prenatal!) you have learned to associate a tone having a harmonic spectrum with a pitch corresponding to the fundamental frequency (or period) of the tone. When you hear a tone, you subconsciously match its spectrum with a template (a spectral pattern) stored in your memory. The pattern that best matches the harmonics in the tone is the pattern that gives the sense of pitch.

An advantage of the template theory is its flexibility. For instance, it can deal successfully with slightly inharmonic tones. Suppose one starts with the 200-Hz complex tone of Fig. 13.3 and shifts each component frequency by 30 Hz. The components then have frequencies 830, 1,030, and 1,230 Hz. This tone does not have harmonic components, and it is not strictly periodic (Fig. 13.5).

Experiments with this tone show that its pitch is not as well defined as the pitch of a periodic tone, but that listeners rather consistently find the pitch to be about 206 Hz. This result provides support for the template theory. According to the

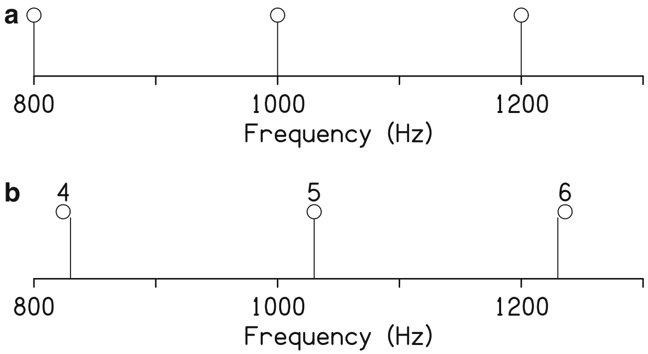

Fig. 13.5 (**a**) The spectrum from Fig. 13.3 is replotted here on a different scale. The fundamental frequency and pitch are 200 Hz. (**b**) The spectrum, shown by *lines*, has components at 830, 1,030, and 1,230 Hz, and these are not harmonics. However, the spectrum fits reasonably well with a harmonic template having a fundamental frequency of 206 Hz and harmonic numbers 4, 5, and 6, shown by *circles*

template model, the listener chooses a pitch for which the harmonics of a harmonic template would best agree with the components that are actually present. In this case, a 206-Hz template would have a fourth harmonic at $4 \times 206 = 824$ Hz, a fifth harmonic at $5 \times 206 = 1,030$ Hz, and a sixth harmonic at $6 \times 206 = 1,236$ Hz. Compared to the actual components, the template fourth harmonic would be too low $(824 < 830)$, the template sixth harmonic would be too high $(1,236 > 1,230)$, and the template fifth harmonic would be exactly right. One can make the case that if one insists on a harmonic template then 206 Hz leads to the best match to the three components actually present. According to the template theory, that is why the brain perceives a pitch of 206 Hz.

Template Model Calculations

To calculate a pitch from a set of spectral frequencies using the template model, you imagine that each frequency in the spectrum suggests a fundamental. You then average all those suggestions. To begin this process, you first need to decide what harmonics shall correspond to the frequencies that you are given.
For instance, given spectral components at 830, 1,030, and 1,230 Hz you might decide that these look like harmonics 4, 5, and 6. Therefore, the suggested fundamentals are 830/4, 1030/5, and 1230/6. The average of these is

$$\frac{1}{3}(830/4 + 1030/5 + 1230/6) = 206 \text{ Hz},$$

in agreement with the more intuitive calculation above. The choice of harmonics 4, 5, and 6 was consistent with the general observation that listeners prefer templates with consecutive harmonics.

13.5 Pitch as an Interpretative Percept

Attempts to build a purely mechanistic theory for pitch perception, like the place theory or the timing theory, frequently encounter problems that point up the advantages of less mechanistic theories, like the template theory. Often, pitch seems to depend on the listener's interpretation. Although the auditory system segregates many of the harmonics of a complex tone like Fig. 13.2 into separate neural channels, the brain finds little benefit in hearing out the different harmonics and tends to hear the tone as a single entity with a single pitch. However, if the spectrum has an unusually large gap, like the gap between harmonics 1 and 10 in Fig. 13.4, the brain decides that there are two entities and assigns two pitches. In this case the place model works.

But the place model does not explain why it is so important for our perception that the partials of a tone should be harmonic. A harmonic tone holds together and is perceived as a single entity with a single pitch. Inharmonic tones, by contrast, often separate into several entities each with a different pitch.

For an inharmonic tone, the brain might hear the complex as a single tone with a pitch for which the harmonic template would best fit the inharmonic components, or it might hear out the individual components themselves. Hearing a single tone is known as *synthetic listening* because the listener synthesizes a single pitch from the components. Hearing the individual components is known as *analytic listening* because the listener perceives the spectral analysis originally performed in the cochlea. Listening to a tone for a long time may cause a listener to hear out individual harmonics—analytic listening. That is why the discussion of the tone in Fig. 13.2 suggested listening for a second or less in order to hear a synthetic pitch. Experiments with ambiguous tones (e.g., Exercise 5 below) show that some listeners tend to be synthetic and others tend to be analytic. The variety of experience in pitch perception seems to preclude any really simple mechanistic model. However, that does not prevent us from continued efforts to find one.

Visual Templates

The interpretative nature of pitch perception has visual analogues, for example Fig. 13.6. You can probably recognize a word in the symbols of that figure, and yet you have never seen the word written like that before at any time in your life. How did you do it? The answer is that the symbols are fairly close to letters for which you have templates stored in your brain, and the template enables you to make the association between the image on the paper and the letter. Also, the word itself is familiar. As you learned to read, you acquired the habit of linking symbols to form a word as a single entity. On the other hand, the last symbol may remind you of an umbrella. Unusual letters can separate like that into several entities, each with its own identity. It's a matter of interpretation, and your brain is engaged in that activity full time, both in the visual world and in the auditory world.

Fig. 13.6 Three (unfamiliar?) symbols can be interpreted as a word or segregated individually

13.6 Absolute Pitch

Absolute pitch (AP), sometimes called “perfect pitch,” refers to the ability to name musical notes without any reference tone. A person with absolute pitch can immediately identify a note played on a piano. Also, that person can sing a note on demand. For instance, if you ask an AP possessor to sing a “D,” that person will sing a note in a comfortable octave that is much closer to “D” than to any other note of the scale. Absolute pitch is different from relative pitch (RP) in which a listener can identify a note when given a reference. For instance, an RP possessor can sing a “D” if he or she knows that the note just played on a piano is the note called “G.”

Absolute pitch in hearing can be compared with absolute color in vision. Everyone with normal color vision has a sense of absolute color in that we can immediately name a color. For AP possessors, musical pitches are identifiable like colors are identifiable for the rest of us.

Less than 1% of the population has AP, and it does not seem possible for adults to learn AP. By contrast, most people with musical skills have RP, and RP can be learned at any time in life. AP is qualitatively different from RP. Because AP tends to run in families, especially musical families, it used to be thought that AP is an inherited characteristic. Most of the modern research, however, indicates that AP is an acquired characteristic, but that it can only be acquired during a brief critical interval in one’s life—a phenomenon known as “imprinting.” Ages 5–6 seem to be the most important.

The musical family tendency can be understood in terms of imprinting. If a child hears songs haphazardly sung in different keys on different days or by different people, the child learns that what is important about tunes is the relative pitches of the tones. If a child listens to a family member repeatedly practicing a piece of music—always in the same key—or if the child regularly practices, the child may learn that what is important is the absolute pitch of tones.

Exercises

Exercise 1, Tonotopic separation

Go back to Chap. 11 and look at the frequency map according to Greenwood showing the place of maximum stimulation on the basilar membrane vs frequency.

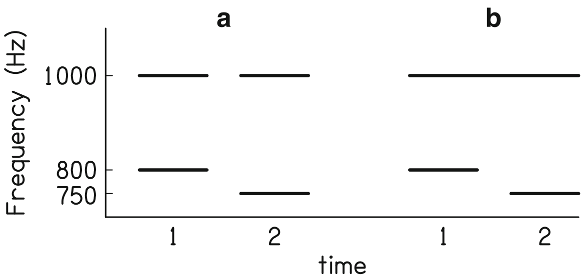

Fig. 13.7 Spectrograms showing the tone pairs in Experiments **a** and **b** called "Analytic or synthetic pitch"

How far apart are places 200 and 2,000 Hz? This large distance is probably the reason that a tone with only these components is not perceptually fused into a single 200-Hz tone. Instead the individual components stand out.

Exercise 2, Neural synchrony: interspike interval histogram

Figure 13.1 Part **a** shows neural spikes in response to an intense 200-Hz sine tone. What time intervals occur between neural spikes? Make a histogram plot that shows the number of spikes separated by one cycle, by two cycles, and by three cycles. When you are done, you will have made an interspike interval histogram—a very useful representation used everyday by neuroscientists. The only difference is that they have thousands of intervals in their data, and you have only 11. Make another histogram for the weak tone in Part **b** of Fig. 13.1.

Exercise 3, A case of the missing fundamental

A complex tone has three components with equal amplitudes and frequencies: 750, 900, and 1,050 Hz. From the template model, what pitch do you expect?

Exercise 4, The template model

(a) Complex tone A has three components with equal amplitudes and frequencies: 420, 620, and 820 Hz. (b) Complex tone B has five components with equal amplitudes and frequencies: 420, 620, 820, 1,020, and 1,220 Hz. From the template model, what pitches do you expect?

Exercise 5, Analytic or synthetic pitch

There is a sequence of two tones, first *1* then *2* as shown in Fig. 13.7.

Tone *1* has two equal-amplitude components at 800 and 1,000 Hz.

Tone *2* has two equal-amplitude components at 750 and 1,000 Hz.

Do you predict that the pitch goes up or down in Experiment **a**? How about Experiment **b**?

♠