S. Perception

This chapter builds on the biology of the previous chapter. The emphasis here is on perception. This topic falls in the area of perceptual psychology. The brain receives the information from the ear by way of the auditory nerve. The sense of sound is perceived by the brain. If we adhere to the strictest definition of sound, which includes external vibrations and internal perception, then there must be a brain for sound to exist. In this viewpoint, the very definition includes the experience of sound.

The Place Theory of Hearing

The *place theory of hearing* states that the place where hair cilia get stimulated along the basilar membrane determines the perceived frequency. This is supported by observation. However, this "place observation" cannot explain some phenomena. Playing a specific tone excites the hair cilia in the corresponding region along the basilar membrane. Playing the tone louder stimulates the hair cilia more and this is perceived as louder.

However, experiments in the 1930s revealed that different pitches carefully matched for loudness were not perceived to be equally loud. In other words, amplitude doesn't solely determine loudness. Pitch influences loudness. We know this from experience since we find high-pitched tones loud and irritating. You might try to defend the place theory of hearing by saying that it is easier to stimulate the stiff region of the basilar membrane. Therefore, high-pitches will of course be perceived to be louder. But maybe this is not the correct analysis. Perhaps other factors and the brain play a role.

There is another observation that presents difficulties if we try to explain everything by the place theory of hearing. Psychologists have found that we can perceive raises in pitch when tones are increased in loudness. We might defend the place theory of hearing by reasoning as follows. If you increase the loudness, the hair cilia shake so much that you get some neighbors shaking also. Since we assumed above that the stiffer end of the basilar membrane responds better, we expect to excite some neighbors at the higher-frequency end more. This raises the tone. But is it enough? Wouldn't the original hair cilia be shaking more wildly and overshadow the higher pitch. These issues are being studied today.

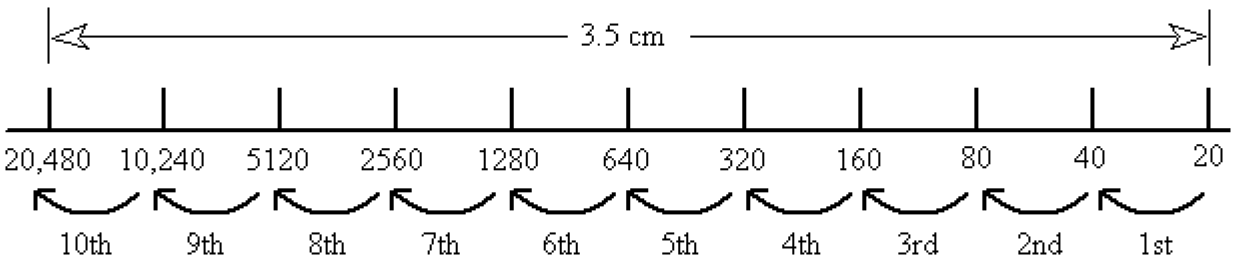
You get the point. We are in gray areas on some of these questions. So although the place theory of hearing is based on the observed spacing of frequency sensitivity along the basilar membrane, it does have limitations in explaining some perceptual phenomena.

Don’t be discouraged by this. Compared to physics, psychology is really hard. The brain and ear are far more complex than compression waves traveling through an elastic medium such as a slinky or air. Also, physics has been around longer, since the 1600s. Psychology as a separate discipline began more recently in the 1800s. We may have to wait centuries before understanding some perceptual subtleties really well.

Let's focus on some simple aspects of the place theory of hearing that we know to be true. Various frequencies are detected along the basilar membrane. Every equal step along the basilar membrane (approaching the oval window) results in doubling the frequency. Therefore, each step corresponds to a pitch increase of an octave. We can count the steps to go from 20 Hz to 20,000 Hz, doubling the frequency each time. We arrive at 10 steps or octaves.

Frequencies along the basilar membrane are seen the Fig. S-1. Most people will have trouble hearing 10 octaves. If we use our more practical range of 30 Hz to 16,000 Hz, we obtain 9 octaves. Note that in either range we use, each doubling step is still an octave. We just start at a different frequency (30 Hz) and make 9 steps instead of 10. But since you need all 10 steps starting from 20 Hz for the entire 3.5-cm length of the basilar membrane, we can say that each octave corresponds to 1/10 this amount, which is 0.35 cm or 3.5 mm.

Fig. S-1. The Ten Octaves of Human Hearing and Positions Along the Basilar Membrane.



The Decibel Scale

We saw how the place theory of hearing gives us a basic understanding of our perception of frequency. Frequency is one of the three fundamental characteristics of periodic waves. The other two are amplitude and timbre. The place theory of hearing also provides us with the essential mechanism for detecting amplitude. The hair cilia, responding at a particular place along the basilar membrane determined by the frequency, get stimulated more when the amplitude of the sound is increased. The greater stimulation of the hair cilia gets sent to the brain through the auditory nerve.

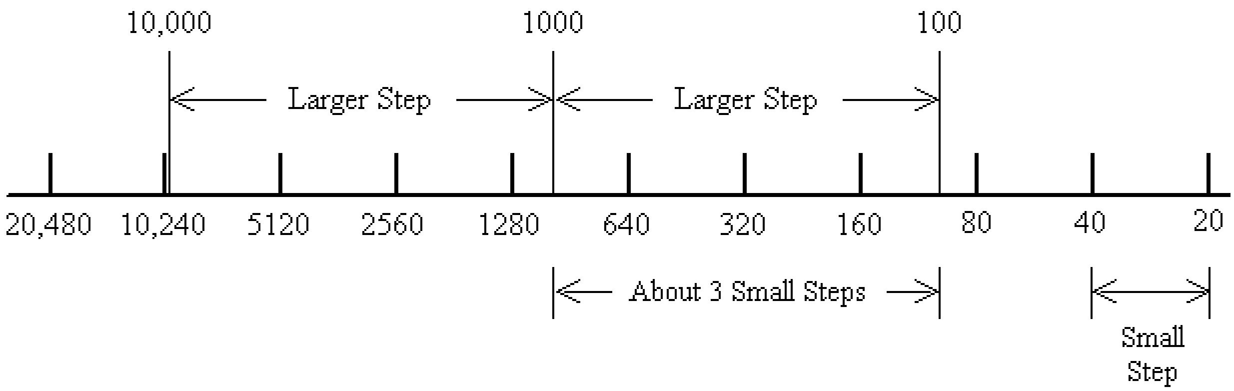
The ear detects a very impressive range of frequencies, from 20 to 20,000 Hz (about 10 octaves). Likewise, the ear detects an impressive range of amplitudes. We can hear the barely audible sound made as a pin drops onto a soft cushion. We can also hear the full blast of an orchestra. We have to be careful that we do not expose ourselves to sounds that are too loud. These can damage our ears. We will take that subject up more fully in our next chapter where we discuss hearing loss.

The way the ear is able to detect so many levels of loudness is due to the fact that it is stubborn in responding to a new level of loudness. The new stimulus must be much greater than the previous one to hear an appreciable increase. We encountered a similar idea in frequency detection. To stimulate the next neighboring group of hair cilia (another 3.5 mm along the basilar membrane), we need to really turn up the frequency. We need to be in the next octave. See Fig. S-2 for another sketch of the frequencies along the basilar membrane. Every time we make a 3.5-mm step along the basilar membrane, the frequency response doubles. In this way, the basilar membrane can pack the tremendous response range of 10 octaves into a total length of only 3.5 cm.

The secret in understanding the strategy at work along the basilar membrane is to realize that equal steps mean you multiply instead of adding. We studied this earlier with the small step size of 3.5 mm. Every small step of this size results in doubling. However, you can take bigger steps. See Fig. S-2 for a larger step size. The secret still applies, but now you multiply by a different number. This number is 10 for the larger step size in Fig. S-2. This "secret" when applied to the psychology of perception is known as *Weber's* (VAY-ber's) *Law*.

The organism will not perceive an equal-step jump in perception unless the original stimulus is multiplied. To get the ear to respond at equal small steps along the basilar membrane, you double the frequency at each step. To get the ear to respond to equal large steps along the basilar membrane, you multiply the frequency by 10 at each step. Note that three of the smaller steps in Fig. S-2 are necessary to make one larger step in Fig. S-2.

Fig. S-2. Weber's Law: Equal Steps Indicate Multiplication.



Weber's Law is usually discussed within the context of loudness. The same idea applies. If you want roughly equally-perceived jumps in loudness, you need to multiply how many sources of the sound you have at each step. In fact, Weber's Law applies in this approximate way to the stimulation of all five senses. Table S-1 lists the application of Weber's Law to the five senses. We employ in Table S-1 the two different step sizes used in Fig. S-2, i.e., multiplying by 2 and multiplying by 10.

There is nothing "sacred" about choosing multiplying factors of 2 and 10 in Table S-1. Weber's Law applies to any number. It just means that the size of the perceived jump will be different. Remember that 3 of the small "doubling" steps equal one of the larger "tenfold-increasing" steps. Refer again to Fig. S-2 to impress this on your memory. This can be also be understood from observing that doubling three times (2 x 2 x 2 = 8) gives us approximately a tenfold increase.

Table. S-1. Weber's (VAY-bers) Law Applied to the Five Senses.

Sense	Small Equally-Perceived Jumps	Large Equally-Perceived Jumps
Sight	Candle Flames: 1, 2, 4, 8, 16.	Candle Flames: 1, 10, 100, 1000.
Hearing	Dropping Pins: 1, 2, 4, 8, 16.	Dropping Pins: 1, 10, 100, 1000.
Taste	Grains of Salt: 1, 2, 4, 8, 16.	Grains of Salt: 1, 10, 100, 1000.
Smell	Drops of Perfume: 1, 2, 4, 8, 16.	Drops of Perfume: 1, 10,100, 1000.
Touch	Strands of Hair: 1, 2, 4, 8, 16.	Strands of Hair: 1, 10, 100, 1000.

The study of stimuli in terms of math and physics is the subfield of perceptual psychology called psychophysics. Fechner (FECK-ner), considered the founder of experimental psychology, came up with a mathematical formula that embodies Weber's Law. Fechner was a physicist and early psychologist. Weber and Fechner both worked in the 1800s during the birth of psychology as a discipline. Weber's Law (or Fechner's mathematical equivalent) is not an exact law; however, it is useful as a starting point in analyzing perception. *Fechner's Law* states that a response is proportional to the logarithm of the stimulus: $R = k \log S$. What's this? Logarithms? Don't worry. You already understand the law if you understand Table S-1. The formula is just the mathematical way of writing the information found in the table.

The area of perception is one of the most challenging applications of mathematics and physics. It is part of the subject of experimental psychology. We are trying to come up with ways, using numbers, to describe how one responds to a stimulus. The detection system offers important clues. Here is where biology and physics come into play. The stiffness of the basilar membrane and its vibrating response to incoming sound is a case in point. Getting a mathematical handle on the perception of stimuli is called *scaling*. Our task now is to scale the perception of loudness. We will present the historical scaling based on Fechner's Law (also Weber's Law). The result is the *decibel scale* we use today as a practical way to measure sound levels. We sidestep working explicitly with Fechner's scaling formula, just as we avoid detailed equations elsewhere in this text. However, it should be stressed that although mathematics is important, it's even more important to understand what's behind the mathematics. What follows is the essence of the historical scaling law for the stimulus-response of loudness.

The sound-level scale is given in Table S-2. It extends the reasoning of Table S-1, where Weber's Law is applied to the five senses. In both tables, we consider dropping pins. The scale numbers are simply counting numbers for the large-sized steps that now continue on for 14 phases. Each perceived jump (step) signifies a tenfold increase in the actual number of pins that drop. The threshold of human hearing is taken to be the sound made by the drop of a pin on a soft cushion. Sound-level meters are designed to get accurate measures of levels. The examples in Table S-2 are approximate. Note the fundamental feature of Weber's Law. The scale rating proceeds by equal jumps or increments while the actual number of sources (pins) for the stimuli goes up by a tenfold increase at each step.

The unit for the scale numbers is the *bel*, named after Alexander Graham Bell. Bell invented the telephone (1876) and made contributions to the study of sound. His mother was deaf and so was his wife. The third column in Table S-2 gives the number of *decibels*. The metric prefix *deci* means one tenth. One tenth of a bel (0.1 bel) is one decibel (1 dB). The decibel is a smaller unit so you need more of them to make up the larger bel. Compared to the reference drop of one pin, a decibel level of 140 is letting 100 trillion pins fall an equivalent distance on the appropriate surface material.

Table. S-2. The Decibel (dB) Scale.

Dropping Pins	Scale	dB	Example
1	0	0	Drop of a Pin
10	1	10	Breathing
100	2	20	Gentle Breeze
1,000	3	30	Whisper
10,000	4	40	Quiet Office
100,000	5	50	Library
1,000,000	6	60	Conversation
10,000,000	7	70	Busy Street
100,000,000	8	80	Factory
1,000,000,000	9	90	Nearby Subway Train
10,000,000,000	10	100	Machine Shop
100,000,000,000	11	110	Construction Site
1,000,000,000,000	12	120	Rock Band
10,000,000,000,000	13	130	Pneumatic Riveter
100,000,000,000,000	14	140	Nearby Jet

The number of pins used to make the sound gives us the intensity from the point of view of physics, not perception. The perceptual scale is the compressed scale that goes from 0 to 14 bels or 0 to 140 dB. We can consider our engineering method of measurement as noting the number of pins we drop to make the sound. So we can consider the number of pins we drop as our intensity. The bel scale is the sound-level scale to approximate our perception of sound. It can quickly be obtained by counting how many zeros there are after the 1 in the number of pins dropped. Therefore, for 100 pins we have 2 (since there are 2 zeros after the 1); for 1000 pins we have 3, and so on.

Finally, to get the decibel column, multiply by the number of bels by 10. This is the prescription given by Fechner's Law: R = k log S. You obtain the stimulus (S) - the number of pins you drop. The "log" is the instruction to count how many zeros are after the 1. Then you multiply by k, which is our multiplier 10. Engineers like to write I_r instead of S. I_r stands for relative intensity; we always compare to the drop of 1 pin. Also, the Greek letter β (beta) is used instead of R (R is used for resistance). We will use β to stand for the sound-level response in dB. Then, Fechner's Law for sound level can be written as $\beta = 10 \log I_r$.

Engineers get precise with the standard of intensity for the drop of one pin. Imagine a drop-of-the-pin sound that is sustained. The energy coming to the area of your eardrum each moment must be equivalent to one trillionth of the energy of a one-watt light bulb falling on an area of one-meter square. This is indeed a small amount of energy. The pin, the cushion, falling height, and distance away all affect the sound level. See Fig. S-3 for a quick overview of approximate sound levels.

Fig. S-3. Overview of Sound Levels (dB).

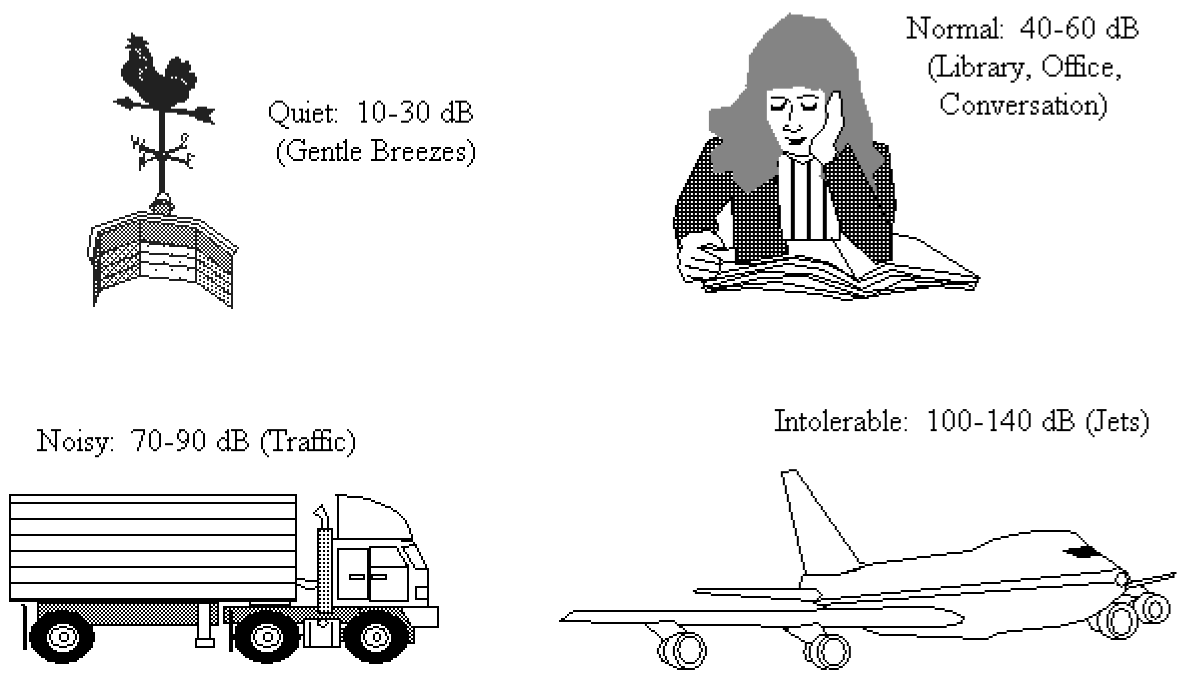


Table S-3 relates sound levels to the language used by composers to indicate on the music score how loud music should be played. These are called dynamic markings. They instruct the performer how softly or loudly to play specific passages and notes. These instructions to performers are traditionally given in Italian. The decibel equivalents given in Table S-3 are approximate. Performers know that sound levels are especially subjective due to an interesting feature of our perceptual process. We perceive a loud sound to be extra loud if it is preceded by silence. So if you have a *ff* passage coming up, play extra softly a little before. Our perceptual dependency on what comes before the sound is just one of the subtleties that makes perception a challenging subject to study.

Table. S-3. Approximate Sound Levels for Musical Dynamic Markings.

dB	Dynamcis	English	Italian
30	*ppp*	Extremely Soft	Pianississimo
40	*pp*	Very Soft	Pianissimo
50	*p*	Soft	Piano
60	*mp*	Moderately Soft	Mezzo Piano
	mf	Moderately Loud	Mezzo Forte
70	*f*	Loud	Forte
80	*ff*	Very Loud	Fortissimo
90	*fff*	Extremely Loud	Fortississimo

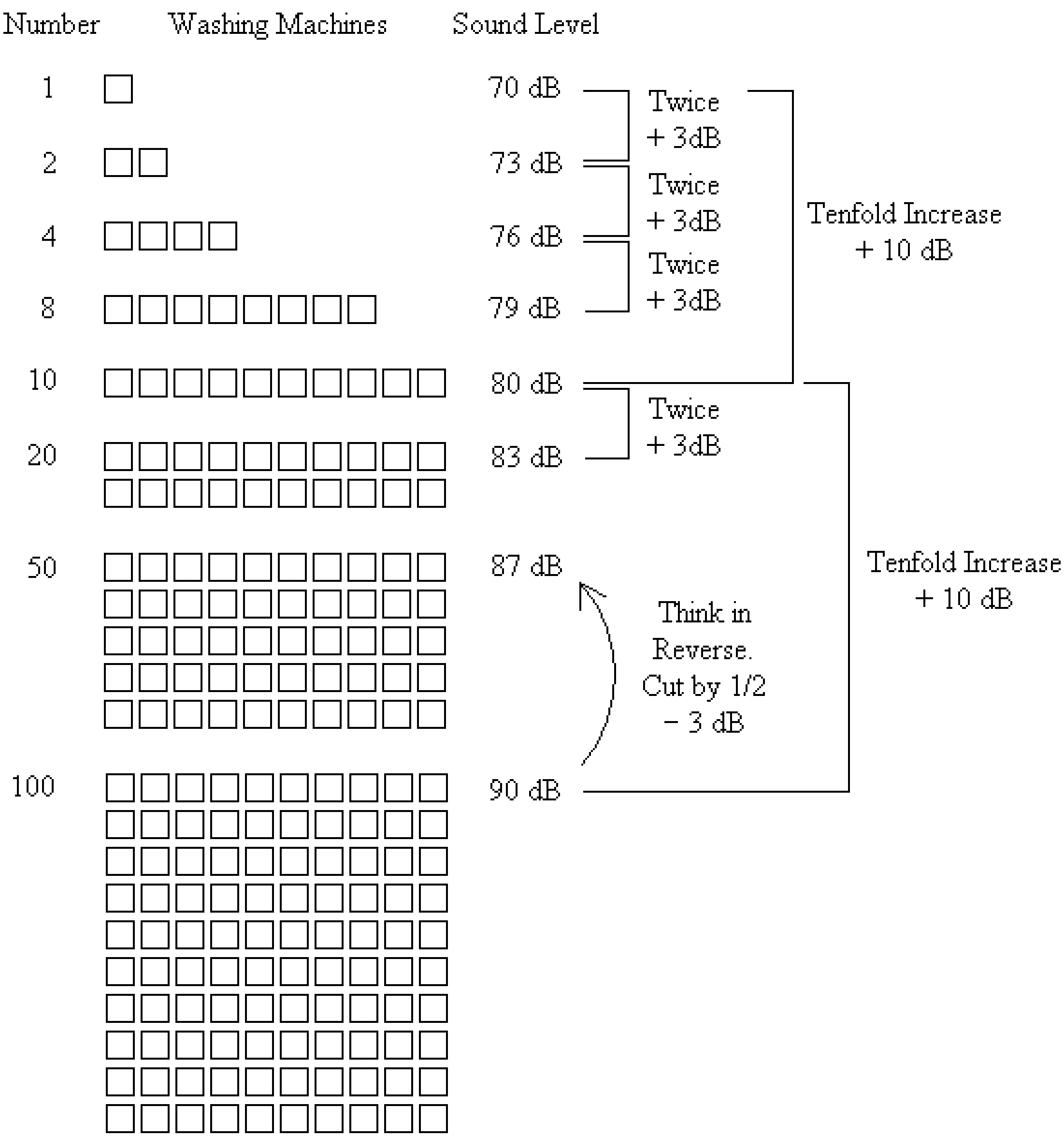
Table S-4 below gives two handy rules for determining sound levels when we increase the number of sources. You may recognize that these rules are the expressions of Weber's Law as applied to the decibel scale. Remember our steps along the basilar membrane. There, one large step (tenfold increase) was equal to about three smaller ones (twofold increases). Doubling the number of sources represents the smaller step size in loudness (add 3 dB), while increasing the number of sources by a factor of ten is our larger step size (add 10 dB).

Table S-4. Two Simple Rules for Determining Sound Level.

Increasing the Number of Sources	Corresponding Change in dB
Two of the original sound sources.	Add 3 to your original decibel level.
Ten of the original sound sources.	Add 10 to your orginal decibel level.

A working example of the rules found in Table S-4 is given below in Fig. S-4. We start with one washing machine at 70 dB. Of course we need to be at the right distance. Assume that we can have more machines at the appropriate distance. Every time you double the amount of machines, you add 3 dB. Every time you multiply the number of machines by 10, you add 10 dB. To get the level for 50 machines, step from 1 machine (70 dB) to 10 machines (80 dB), then to 100 machines (90 dB), and cut the final number of machines in half. You then subtract 3 dB instead of adding. With our two rules, you can determine so many cases. You can easily estimate the levels for other amounts in between. For example, 7 machines produces about 78 dB. Why would 5 machines give 77 dB? Do not make the common careless mistake and state that 2 machines would be 2 x 70 dB = 140 dB. Note that 100 machines produce only 90 dB.

Fig. S-4. Example in Determining Sound Levels.



The study of how frequency relates to loudness was undertaken in the 1930s. The reference for the sound-level scale (dB) is a 1000-Hz tone. Think of a precise experiment where we do not drop pins but use flutes that produce a barely audible 1000-Hz tone at a given distance. Then, at the proper distance, 1 such flute gives 0 dB, 10 flutes give 10 dB, 100 flutes give 20 dB, and so on. If we work with another frequency, subjects perceive a different "loudness spectrum."

So we modify our table from the very early chapter concerning the physical and perceptual characteristics of sound. This table is reproduced in Table S-5. However, we add a qualifier to emphasize that the correspondence is approximate.

Table S-5. Approximate Relationships Between Physics and Psychology.

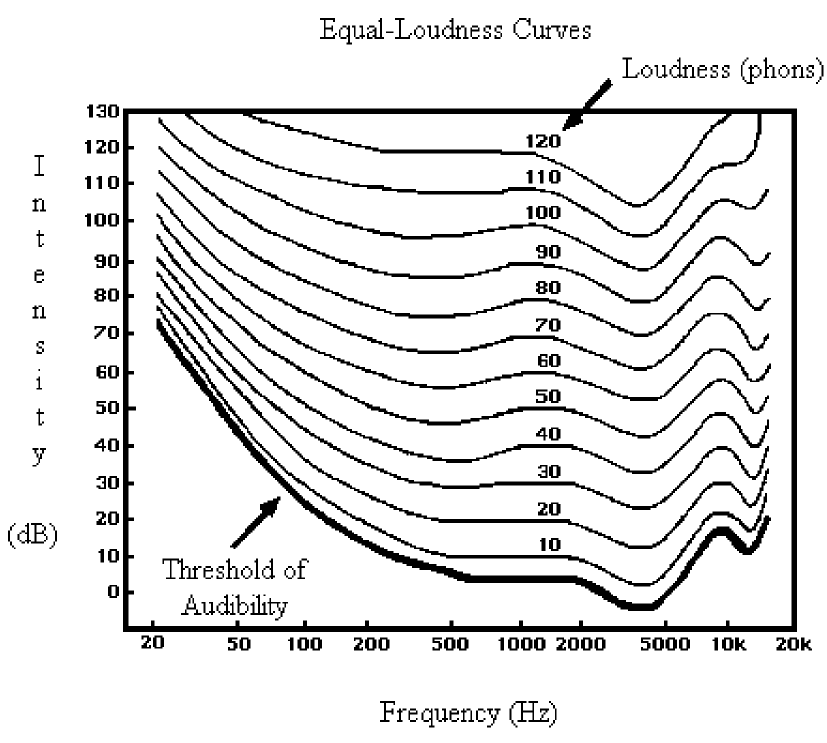
Physical	Perceptual
Amplitude	Loudness
Frequency	Pitch
Waveform	Timbre

Fletcher and Munson (1933) made a study of the perception of equal loudness and how the sensitivity of the ear varies across the frequency spectrum. They started with the sound-level scale which assumes a 1000-Hz tone. They then presented subjects with different pitches. Imagine replacing the 1000-Hz flute with one at 500 Hz. We then play one of these, then 10, then 100, and so one. Of course, in practice one uses a tone generator and controls the energy output to simulate the series of cases, 1, 10, 100, 1000, and so on. Fletcher and Munson found that their subjects perceived different loudness levels for the different frequencies played at the same level according to a scientific instrument. For example, if we employ the 1000-Hz flutes, we get a threshold response when one such flute is played. Now if we switch to a 50-Hz instrument (bass tone), we need 10,000 instruments to just get the subject to hear anything. This is 40 dB higher! We are less sensitive to bass tones than we are to 1000 Hz.

The Fletcher-Munson experiment carefully starts with a set pitch. We push up the decibel level, monitoring it on a scientific instrument, until the subject hears something. This establishes the threshold for the pitch. We then draw a curve across the audio spectrum which represents thresholds (see the lowest curve in Fig. S-5).

Other equal loudness curves are determined using 1000 Hz as the reference. By definition, the perceiver is in agreement with the sound-level meter at 1000 Hz. This phase of the experiment begins with the 1000-Hz reference tone along with the pitch other than 1000 Hz. The reference is set to a decibel level according to the sound-level meter. For example, the 1000-Hz tone might be set to 30 dB. Then, the other tone with a different frequency is presented to the hearer. The hearer adjusts its volume so that the different frequency matches the loudness of the fixed 30-dB reference of 1000 Hz. Tones that sound the same in loudness, are found to have different decibel levels according to the meter. In our above example, 40 dB at 50 Hz sounds as loud as 0 dB at 1000 Hz. To avoid confusion, it is said that they have the same value in *phons*, the subjective scale.

Fig. S-5. Fletcher-Munson Curves.



Find any point along any curve in Fig. S-5 as follows. First choose a specific curve, then a point along that curve. Suppose you choose the 30-phon curve. You then slide along this curve to any point. Consider stopping at the point to the left, corresponding to 50 Hz (horizontal) and 60 dB (vertical). This point tells us that in order to hear a tone of 50 Hz at the same level as the reference 1000-Hz tone at 30 dB, we need to make the 50-Hz tone 60 dB. In other words, 50 Hz at 60 dB has the same loudness as 1000 Hz at 30 dB. Each is said to have 30 phons. Note that the dB-value and phon-value agree at the 1000-Hz reference for all levels of intensity.

As we move to the outer limits of human hearing, the curves rise. Focus on the lowest curve, the threshold curve. This curve describes barely audible sounds. The threshold curve gets higher at each end of the spectrum. Note the enhanced sensitivity near 3000 Hz. Here all the curves dip down. The ear canal is like a small pipe and has a resonance frequency near 3000 Hz. The ear canal amplifies sound near 3000 Hz as a resonance effect. The threshold for low frequencies is high. This difficulty in hearing low bass tones is actually good. Otherwise, we would hear the low-frequency sounds made inside our bodies. Since the ear is not very sensitive in the low-frequency range, sound-level meters have a special weighting mode (A-weighted) that discounts lower frequencies. Meters also usually have fast and slow response modes, the slow response giving more or less an averaged sound level.

Other Perceptual Phenomena

1. Masking.

When more than one sound is perceived, the louder sounds are heard more easily. Therefore, it is possible for a loud source to overpower a soft one. This can happen to the point where we no longer hear the soft one. This is called *masking*. We all have experienced the trouble of hearing something soft because something else is louder and distracting.

White noise can help mask sounds. White noise presents us with all frequencies. We have encountered such examples as the fan, sound of the ocean, and rain. Putting on a fan helps some people go to sleep due to the masking effect. Distracting sounds are covered up by the soothing even-distribution of all frequencies.

2. Periodicity Pitch.

Masking is an example where loud sounds prevent us from hearing other sounds. Here we see that certain sounds can cause us to hear other sounds not originally present in the source. The place theory of hearing cannot explain why we perceive tones that are not present in the sound waves. Consider sending a 200-Hz sine wave to the ear along with a 300-Hz tone. The brain perceives the 200-Hz and 300-Hz tones. It recognizes that these two tones can be thought of as the second (H2) and third (Hz) harmonics relative a 100-Hz sine wave. The brain registers at a lower level of intensity the fundamental at 100 Hz (H1), the *periodicity pitch*. A frequency-analysis of the incoming waves is done by the ear-brain system, establishing and perceiving the fundamental!

We saw that each periodic tone can be represented by a Fourier spectrum of harmonics. Most of the time the fundamental is the strongest component. The ear-brain expects the fundamental to be there and puts it in if it's not. This explains why we hear bass better than we should from a small 2-inch speaker. The low fundamental tones are lost to some extent, but the ear-brain system supplies them. The ear-brain knows they should be there.

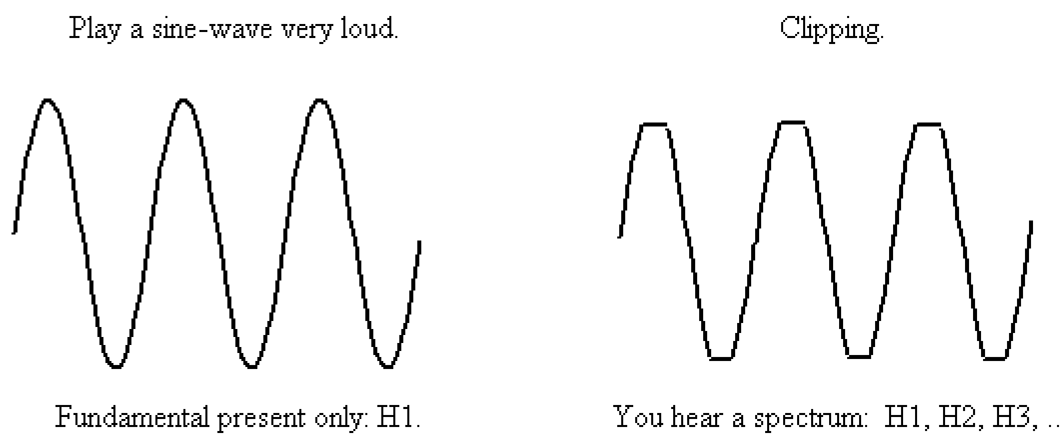
Now consider a rephrase of our earlier question about a tree falling in a forest. Is "sound" present for a fundamental tone if it is heard but there is no source making that frequency?

A *periodicity theory of hearing* has been developed based on observations that we hear fundamental tones not present in the incoming sound. In such a theory, Fourier analysis of incoming waves are relevant. Both the place theory of hearing and the periodicity theory of hearing are important in providing for a more complete picture of hearing.

3. Aural Harmonics

Another fascinating case of hearing components of sound not present in the original sound is experiencing aural harmonics. A sine wave has one harmonic, the fundamental (H1). However, if you play it loud enough, the eardrum can't vibrate through the distance it needs to. You get clipping of the wave as the eardrum reaches its limits. The information sent to the middle and inner ear is now no longer a sine wave. Therefore, you perceive overtones (*aural harmonics*), frequency components not in the original sound entering the ear. See Fig. S-6

Fig. S-6. Aural Harmonics Due to Clipping.



The amplitude of the sine wave is so great in Fig. S-6 that the eardrum cannot faithfully reproduce it. The wave gets distorted. If we turn up the volume too high, the system is not able to handle it. Distortion can occur in making tapes if we tape the source with the amplifier setting too high. The strength of the amplified sound is indicated in recording equipment by a sound-level display to guide us. Whenever an electronic component distorts the shape of a sine wave, harmonics of the sine wave appear. This is referred to as *harmonic distortion*.

4. Combination Tones.

Playing two very loud sine waves causes us to hear additional tones beyond those discussed above. We hear the original tones, the sum and difference tones at low levels, and possibly even more tones. For example, if we play a 500-Hz tone very loud with a 700-Hz tone, we hear 1200-Hz (sum) and 200 Hz (difference). The lower tone may be difficult to hear due to our lack of sensitivity to low pitch.

The resulting frequencies are the *combination tones* made by combining the original frequencies. Other combinations found by adding and subtracting various multiples of the original frequencies may also be heard. How can you use beats to determine if a 1200-Hz tone is heard when a 500-Hz tone is played loudly with a 700-Hz tone?

5. Binaural Effects.

Binaural effects are phenomena that result from our having two ears. Just as two eyes (binocular vision) give us an excellent sense of three dimensions, two ears provide us with a better three-dimensional sense of hearing. With two ears, we can more easily tell from which direction a sound comes. Sounds at our left do not reach the right ear as well. The brain constantly compares the sound level at each ear to give us a perception of our surroundings. For long-wavelength bass tones, the brain relies more on a comparison of phases. When a compression reaches the closer ear, there is a delay before the compression reaches the farther ear due to the extra distance. So different parts of the wave cycle reach each ear at any given time.

The important role of the brain in processing signals from the auditory nerve is evident in the following experiment using two ears. Earphones are employed to send a different signal into each ear. When the different tones are close in frequency, we hear beats. Even when care is taken to play the tones softly to eliminate any bone conduction in the skull, beats are still perceived. The conclusion is that the beats occur in the brain. When we usually hear beats, the waves combine physically outside the ear. The pressure waves add. The result is a fluctuation in the sound wave itself. You can hear it with one ear.

When the tones are separated, they cannot physically add together. Each tone enters a different ear. However, the combination of these signals from the auditory nerve of each ear to the brain is processed in the brain. The brain effectively combines the waves in a way similar to the physical addition of wave amplitudes. The beats experienced are perceptual or psychological rather than physical. Do they really exist? Is there sound if a tree ... ?

--- End of Chapter S ---

