CHAPTER 3

FROM VIBRATION TO SENSATION

How is sound created and how does it reach our ears? Using accessible notions of physics, anatomy, and neurophysiology, this chapter will explore sound from its various sources of energy, how it is transferred through different media, and how we can categorize different types of sounds by their basic characteristics of rhythm, intensity, pitch, timbre, speed, shape, and organization. The construction and function of the ear as receptor to these sounds is then introduced, so that we may understand what the effects of sound have on our physical body.

The source

Energy. Vibration. Our universe defines itself with the interaction of masses and forces, both visible and invisible, physical and emotional. The point of origin may be two galaxies colliding, a drop of rain in a bucket, or a child's wail of hunger. They all produce vibration, some reaching our auditory sensitivities, and some able to attract our attention more than others. They all contain motion, a dynamic seeking equilibrium, a goal. In the search for this equilibrium some will produce a wave of energy that sets into motion air molecules at certain frequencies and loudness, enabling us to perceive sound.

Friction between air molecules and solids can give rise to whistles, wind, and sounds from woodwind instruments. With liquid, the air creates gurgles, gargles, and busted bubbles. Water with water makes splashes and waves, and with solids makes cannonball dives, squirt guns, and garbage disposals. Solids against solids range from the lowest earthquake rumble of tectonic plates to the fizzy brush stroke on a cymbal.

Exploring the possible sources of sound-making is one of the most joyous processes for the sound designer. Tuning into what we already hear around us and separating the sources is the essential ability needed for creating new sounds and combinations. The **Sensitizing to sound quality** experimentation later in this chapter (see page 69, **TRY THIS**) will help you pay attention to the possibilities of the world of sound and to your talent to sense the vibrations.

The medium

Between the sound source and your ear is a space across which the vibrations must travel. This medium can be solid, liquid, or gaseous — usually our atmosphere transmits the sound at about 600 feet per second. (Did you ever notice the crack of the bat at a big ball game arriving a bit delayed from the hit, or the wait for thunder after lightning depending on how far away it is?) Due to the varying elastic, restorative properties of the different states, sound waves travel through water four times faster than through air, and even faster through solids. You can sense this phenomenon by tapping a 20-story metal stairwell with a friend on the other end who will feel the metal vibrate before the arrival of the air-transmitted sound.

While air molecules vibrate easily as sound waves pass through, water molecules are denser and slow down with friction between each other, allowing much less sound to be transmitted. An excellent example of this contrast between air and water occurs in the opening sequence of *Saving Private Ryan*. Sound designer Gary Rydstrom created the indelible sensation of a gun battle raging from the cliffs of Normandy down to where the soldiers in waist deep surf are being shot down. When the camera submerges, the sound environment transforms radically with the eerie isolation of underwater events, an illusory haven from the chaos above.

Why then would you hear a distant motorboat engine more loudly underwater than above? The surface is reflecting the sound energy from the propellers back into the water, with little escaping into the air, much as an optic fiber maintains the light beam intensity until it reaches the receptor.

Unlike light, sound can travel through solid materials like walls, although the high frequencies are usually filtered more than in the gaseous medium. Sound also bends around corners more easily than light, giving less privacy than our visual world. This property lends itself to telling much more of the film story than what is seen on the screen (see Chapter 7, **Sound and Image**).

Sound qualities

To identify, manipulate, and create audio in a system of **sound-plus-listener**, it is very useful to be able to label the types of sound by their qualities, which also facilitates communication between creative members on the film crew. The major attributes of sound fall into the categories of **rhythm**, **intensity**, **pitch**, **timbre**, **speed**, **shape**, and **organization**. Our ability to perceive these parameters and their extremes is governed by the physiologic limitations of the hearing apparatus, which we will discuss later.

TABLE 3-1
CATEGORIES OF SOUND QUALITIES

Quality	Extremes
rhythm	rhythmic-irregular
intensity	soft-loud
pitch	low-high
timbre	tonal-noisy
speed	slow-fast
shape	impulsive-reverberant
organization	ordered-chaotic

Rhythm

Rhythm, or the lack of it, characterizes sound through time, ranging from an absolutely regular clock tick or resting heartbeat to the spastic squeals of feeding pigs or the cacophony of a bicycle crash. Organic sounds can be either **rhythmic** (breathing, brushing teeth, woodpecker) or **irregular** (conversational speech, whale songs, volleyball game). Mechanical sound tends toward predictability, until it signals some kind of breakdown or outside intervention. The predictability of a sound can lend a certain tranquility and assuredness, or nagging oppression. An irregular sound can keep you alert, frightened, confused, or in fits of laughter.

Intensity

The **loud** to **soft** continuum is measured in energy increments called decibels, a logarithmic scale of sound energy with each ten points representing ten times the loudness. True silence is perhaps impossible to attain. At the extreme of zero dB (decibels), considered the bottom threshold of hearing, even the vibrations capable of moving the eardrum only a fraction of the diameter of one hydrogen atom can produce a measurable sensation. Our ear mechanism creates its own noise as well, so absolute silence is virtually impossible. Silence therefore must be tailored for dramatic needs, as will be discussed later. At the other extreme, sustained loud sounds can damage hearing, as may explosive attacks, and they cause fatigue in the listener if overused.

TABLE 3-2 INTENSITY OF SOUND

dB	Units of Energy	Examples of Sounds
0	1	threshold of hearing
10	10	rustling leaves
30	1,000	whisper
40	10,000	quiet home, city background
60	1,000,000	ordinary conversation
70	10,000,000	rush-hour traffic, average film dialogue level
80	100,000,000	loudest TV sounds
90	1,000,000,000	loudest peaks in analog theatrical films
100	10,000,000,000	shouting, jackhammers, motorcycles
110	100,000,000,000	loud rock music, loudest peaks in digital theatrical films
120	1,000,000,000,000	jet take-off
130	10,000,000,000,000	threshold of pain

Although our hearing extends over an enormous 1 to 10 trillion relative range, the actual energy of the human voice, for example, has only the power equivalent to one millionth of an ordinary light bulb.

According to "The Sonic Boom," an article published by *Playboy* in May, 1967, the loudest continuous sound ever heard on earth was a siren at

175dB, which could make pennies dance in a vertical position and burst a cotton wad into flames within six seconds! Of course film sound will never get near to this level, but the power of sound can be awesome indeed.

Pitch

The parameters of pitch follow the gradient of **low** to **high** frequencies, in normal hearing ranging from about 20–20,000Hz (Hertz, or cycles per second), although older people's upper range is often decreased. Very low pitches (known as infrasonic) are felt bodily as rumblings more than acoustic phenomena, since the frequency gets so slow that individual beats become distinguishable as rhythm rather than pitch. Frequencies above hearing range (known as ultrasonic) may not be audible but can cause uneasiness if emitted loudly.

TABLE 3-3 PITCHES OF SOUNDS

Hz	Examples of Sounds
10	earthquake
20	lowest range of hearing
27	lowest note on piano
50	low range of singing voice
80	low range of male speaking voice
263	middle note on piano
400	high range of child's or woman's speaking voice
1,000	high range of singing voice (fundamental tone)
4,186	highest note on piano
10,000	hiss of spoken consonants (s, ch, z, f, th)
20,000	highest range of hearing

Timbre

When sound waves pulse at regular intervals (also referred to as periodic), they create a pure or **tonal** sound, as opposed to a **noisy** sound made of overlapping and intermingling frequencies which produce highly complex waveforms (nonperiodic, irregular intervals). In the continuum between a pure tone and noise lie musical instruments that emit several frequencies known as **harmonics**. The harmonics for each instrument have regular waveform patterns, defined as timbre (pronounced *tam'-ber*), ranging from the simple harmonics of the flute or triangle to the complex pattern of the bassoon, violin, or timpani. The human voice has a tonal quality when it is singing, or more of a noisy quality when it is coughing or sneezing. At the noisy extreme we find explosions, wind, splashing, clapping, and what is referred to as "white noise," which is electronically generated with a completely mixed, noncharacteristic frequency range.

(Distinction should be made between this concept of a *noisy* sound with another definition of **noise** that refers to any undesired sound signal that impairs transmission of an intended message, be it musical, verbal, or otherwise.)

Speed

When acoustic impulses are repeated, they can fall between the extremes of **slow** and **fast**. When a sound slows down beyond our conscious ability to perceive it as continuous, with a pause longer than about a second, our attention becomes distracted. However, with additional acoustic cues such as a melody or verbal context, we can integrate information at a still lower rate. Resting cardiac pulse and the lethargic march of a funeral procession are examples of slow forms of soundmaking. At the upper extreme of 20 beats per second, the individual sounds begin to blur into a steady pitch (or low frequency). When speaking, the highest optimal speed for comprehension is about five syllables per second, with frequent interruptions to give the listener time to process the information.

Shape

Another technical term for this parameter is the **envelope** of sound, defined by its attack (onset, growth), body (steady-state, duration) and decay (fall-off, termination). The gradient of measurement of shape ranges from more **impulsive** to more **reverberant**. An explosive gunshot in an open area with no echoes would be termed as impulsive, beginning rapidly, peaking, and decaying rapidly. On the opposite extreme would be the reverberant sound of wind howling through a tunnel, gradually rising and falling. The listener's perception of the sound shape depends not only on the waveform created by the source, but also on the distance and reverberation properties of the surrounding space. Note that an echo differs from reverberation in that the former is a repetition or partial repetition of a sound, whereas the latter has no distinguishable repetitions.

Because every natural sound has a beginning, or attack, there will never be a mathematically pure tone or sine wave in a normal environment. The striking of a tuning fork, for example, sets up little imperfections called **onset transient distortions**, telling our ears that it is a tuning fork as opposed to an electronic oscillator.

TRY THIS

Sensitizing to sound quality

To fluently communicate with all these variables in sound, you should be very much in touch with the experience of hearing them. This is not a common form of observing. Set aside a moment and sit quietly in your home, a park, or a restaurant, and count how many different sounds make up the sonic environment. Look for sounds with particular qualities, for example, high-pitched, organized, or fast. What combinations do you hear? If there are sounds that are similar, which aspects are similar and which are different? Notice that the more the similarities, the harder it may be to distinguish between them. How does it feel to inspect every sound carefully during your search? This listening game hones the skills of a good sound designer, and you might find a simple joy in these activities — especially when they lead to creative discoveries for your film work.

At another moment after you have tape-recorded some conversation or group discussion, listen to the tape and concentrate on the sounds that were not intended to be recorded. What background noises, interruptions, or surprises appeared? How would you characterize them for their sound qualities, and what influence do these qualities have on the intelligibility of conversation or on the concentration of the participants? Later we will see how perception principles of the ear and brain relate to these questions of sound quality.

Organization

This quality pertains to how orderly the acoustic signals are to the human ear, ranging from **organized** to **chaotic**. Unlike rhythm, which derives from physical and biological parameters, organization depends greatly on the listener's social and educational background as well. A foreign language, for example, will seem utterly unintelligible and chaotic to the unacquainted until the meaning is learned, and then it translates into organized concepts. This can also apply to music, ambience, or effects, where organization delights in meaningful rhythm, tone, intensity, etc., while chaos reigns in cacophony, dissonance, and disarray.

Physical effects of sound

Sound energy can influence matter directly and be seen in physical forms and shapes of intricate geometric figures. Vibrations that pass through a drumhead with sand or through a basin of water can produce oscillating images of either beauty or chaos, depending on the organizational quality of the sound. A low *omm* chant, for example, produces perfect concentric circles that change to wobbly edges when a high *eee* is sounded.

The force of sound can be destructive as well as creative. Through the principles of resonance (see ***Entrainment***, Chapter 4) a strong soprano singer can shatter a glass with her voice and a lithotripter can break up kidney stones by using explosions of sound at the same frequency as the resonant frequency of the stone mass. These extremes are not often reached in film sound, but it is useful to know where your limits and potential lie, especially with respect to how you may be physically affecting the human body and mind.

As a general rule, the lower frequencies up to around 65Hz will resonate in the lower back region, pelvis, thighs, and legs. The timpani, or orchestral kettledrums, are a prime example of a sound that activates this region not through the ears, but directly — affecting sexual, digestive, and deep-seated emotional centers. As the frequencies increase, effects are felt more in the upper chest, neck, and head, influencing the higher biological functions of the nervous system and mind.

Sound can affect our body temperature, blood circulation, pulse rate, breathing, and sweating. Loud music with a strong beat can raise body heat, while soft, floating, or detached, abstract music can lower it. This phenomenon can be used in sound design to accentuate or counterpoint a scene with, for example, a cold winter day or a searing rocky desert.

Noises can energize, release pain, and dissipate tension, in particular those from our own voices. Helping break through to new levels of achievement, focused energy is emitted with the strong vocalization of a karate punch — *hai!* But a sound can also bring about negative changes, as with obnoxiously loud factories, jackhammers, or jet planes. A buzzing saw close to the ear can bring immediate headaches and extreme disequilibrium, and loud low frequencies can create stress and internal pains. Drones produced by air conditioners or the last days of a fluorescent light bulb that go on for hours can have an unbalancing effect, but the drone from a richly harmonic organ or your own voice can stir wonderful emotions and break down blockages in the body.

TRY THIS

Helen Keller, the college-educated blind and deaf woman, "heard" sound through touch and vibration. You can gain a valuable experience in replicating Helen's heightened sensitivity by temporarily removing your sight and hearing with a blindfold and earplugs. Select an environment that will not make you self-conscious. Begin projecting a tone with your own strongest voice, first in the lower registers, and place your hands on various parts of your body — abdomen, chest, throat, forehead. Notice where the sound vibrates on your hands. Raise the tone to normal speaking pitch, feel the vibrations throughout the body, then raise the pitch as high as you can, very loud. Does the sound move up your body as well? Try this now with a loud speaker that has a strong bass, playing a selection of music that isolates low and high frequencies at different moments. Move your body very close to the speaker and see which tones resonate with which areas of the body. By sensitizing to this, you will be conscious of an effect that for the most part will be unconsciously affecting the film audience.

The ear

"Our ears are open before we are born. Our consciousness begins with them. Is that the real reason why we can never, ever close our ears so long as we live?"
—Joachim-Ernst Berendt

Hearing development

Hearing is the first sense that develops in the womb. By the fifth month of pregnancy, the cochlea of the fetus's ear is fully developed and sounds from both inside and outside the mother's body are perceived. Someone asleep can accelerate their learning from instructional language tapes. When a patient is anesthetized or in a coma, the auditory fibers continue to transmit information to the brain that is later consciously recalled. It is said that hearing is the last sense to leave before death, and Tibetans believe that listening continues beyond this moment into the hereafter.

Nevertheless we often demote the ear's role to that of an auxiliary organ, almost a reflex in the same way that it maintains our physical equilibrium automatically. We become conscious of the ear's function only when the information provided by the eyes becomes totally inadequate. In a way, though, this gives the sound designer great power to work behind the scenes of the audience's attention, directly into their subconscious.

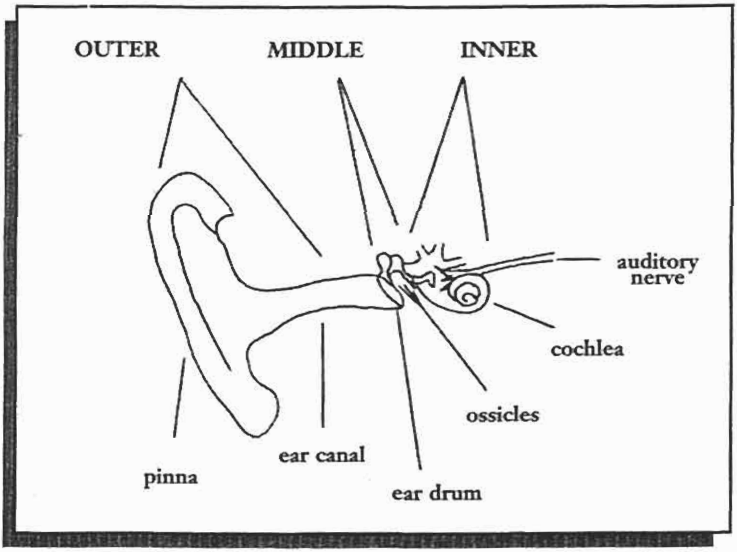

Figure 3-1 cross-section of ear (not to scale)

The function of the ear is intimately related to its structure, as are the sounds entering the ear related to the sensations we perceive. The three parts of the ear are outer (pinna/ear canal), middle (eardrum/ossicles), and inner (cochlea).

TRY THIS

Standing near a running shower, constant traffic, or radio static, gently place your fingers over both pinnae (plural of pinna) without blocking the entrance to the ear canals. As you alternate between covering and uncovering your pinnae, notice how the high frequency hissing sounds are influenced, but the bass tones remain constant. Notice also some loss of directionality when you turn your head with your pinnae still covered. Listening to the "ocean" in a conch shell is a related phenomenon in which the shell amplifies certain frequencies in the ambience to create the effect of a distant ocean roar.

Outer ear (pinna/ear canal)

The rubbery mass of folds and ridges we see on the side of our heads and what we generally call the "ear" is in fact just part of the outer ear. The pinna's main job is to amplify the ambient sounds and funnel them into the ear canal. It also accentuates middle-high frequencies between its folds, resonating in the range of the human voice, and helps us locate sound sources in space, giving minute clues through the varying resonance qualities from different sources. The sound then travels down the ear canal toward the middle ear.

Middle ear (ear drum/ossicles)

The sound energy next vibrates the eardrum, which is attached to three tiny hinged bones called the ossicles. Why the complexity? The ear is transforming sound energy from the medium of air to the medium of water in the inner ear, and requires an amplification of this energy to move the sluggish water molecules. Like the pinna, the ossicles boost the middle-upper frequencies of the voice over other sorts of sound. They also serve as a kind of pad or limiter for very loud sounds, with tiny muscles beginning to respond within 1/100th of a second of the onset of a dangerous sound. But these muscles can't react quickly enough to an explosive gunshot and will tire from constant assault, so they are not perfect in the protection against deafness. Our own voice sounds so different to us when we listen to it recorded, because we usually hear both by way of our ear canal and through direct vibration of the ossicles from the bones of our skull.

Inner ear (cochlea)

Finally, as the sound vibrations enter the wet inner ear, the energy is converted into information that the brain can use. Up until this moment the sound has been processed, transforming from the mechanical movement of air molecules to water molecules. Now the sound will be sensed and become information, with the mechanical energy being converted into neuronal energy in the **cochlea**, a kind of flesh-and-blood microphone. Inside the conch-shaped cochlea the vibrations agitate tiny hair cells like grass waving in the wind. Hair cells of different lengths correspond to different sound wavelengths, longer hairs for lower sounds, shorter for higher.

Curiously, there are a mere 14,000 sound receptor cells in an ear, compared to 1 million visual receptor cells in an eye. But if you play two pure pitches simultaneously you can distinguish them, while if you project or mix two pure colors together, they blend into one solid color.

Sensitivity

Pitch, intensity, and speed are sound qualities that the ear, because of its inherent physiology, will respond to differently than will an electronic measuring device. There is constant interaction between these parameters as far as our perception is concerned. For instance, intensity can influence speed perceptions (a loud tone will sound longer than a soft one), pitch will affect intensity perceptions (a high tone will sound louder than a low one of the same measured volume), and duration will affect intensity (a tone of the same loudness will appear to grow weaker over time).

Because of the physiology of the ear, the midrange frequencies are accentuated, helping us distinguish voice information in a mixed frequency environment. We can also sleep more soundly without the distraction of our internal bodily noises like heartbeat and breathing. So when you need more bass, you have to account for the extra energy to make it be heard, which may risk saturating the film's optical soundtrack. An example is the basso profundo voice of Jabba the Hut in *Return of the Jedi*, needing tremendous intensity to sound normal to our ears.

At lower volumes our ears are even less sensitive to low and high frequencies, so when music is played softly, the loudness button on the stereo will boost these frequencies to better match the apparent level of the middle range.

When two tones are heard within a frequency **critical band** of approximately one-third of an octave, the increase in perceived intensity is not so much as when the two tones fall outside this critical band. So a way to make a sound seem louder is to increase the spectrum of frequencies. If, for example, you want a rumble to be more present, then add middle and higher frequencies, not more lows.

The speed of the tone by itself or as a sequence can influence our perception. When it is singular, the loudness of the tone will take about a third of a second to register fully (called the **integration time**), so less than this will lower its perceived intensity. In a sequence of loud sounds, the later bursts will seem softer (termed **aural reflex**). This is fundamental to observe in big action movies: Avoid constant sound blasting, because the effect will be diminished with no aural relaxation for the audience. Contrast is essential for this and many other parameters to work on a dramatic level.

Masking

A specific phenomenon of diminished perception of sound is explained by masking. This can occur because of frequency or temporal causes. In general, our ears will not be able to hear softer sounds under a louder sound of the same pitch, which is called **frequency masking**. When editing production tracks this is extremely useful to solve the problem of background noises jumping from one cut to another. Placing over the entire scene another ambience track of the same frequency range as the production track, but a little bit louder, will smooth out all the cuts. For medium frequencies, the masking occurs when the signal is at least two and half times louder than the background noise (signal-to-noise ratio greater than 1:2.5 or at least 4dB).

In temporal masking, a loud sound will cover up a softer sound that follows a short time after. This happens because the auditory system has clamped down from the initial high volume and this is termed **forward masking**. But because the brain responds to loud sounds more quickly than to soft ones, **backward masking** can occur in the opposite direction. These physiological effects can help the sound editor cover up unwanted discontinuities, but also must be accounted for to avoid unwanted loss of auditory information.

Masking has a stronger effect when the two sounds come from the same direction. This can be taken care of in a pre-mix so that the sounds are on a single track, or the sounds can be kept separately and assigned to the same track in the final mix. In the latter case, remember to note future track assignments. Otherwise, the signals will need to have a greater ratio to affect the masking.

Phase shifting between two signals (when their sound waves do not strike the microphone at the same part of their cycle, therefore causing some cancellation or interference of the sonic energy) will emphasize their distinct sources. So if masking is desirable, the phases should be in sync. On the other hand, two similar sounds might be played with, to separate them intentionally using the phase-shifting technique. This is normally more detectable when listening with headphones to isolated sounds in each ear, so it would be more applicable to the new media on computers than to theatrical film exhibition.