Chapter 1
A Point of View

If you had to choose, would you prefer to be deaf or blind? Without taking time to reflect, most of us would prefer to be deaf, to see rather than hear. Now imagine going about your day-to-day life deaf or blind. Blind individuals can make it. They communicate, laugh, and joke with blind and sighted individuals; they listen to music and so on. Deaf individuals make it only with extreme difficulty and are generally restricted to the friendship of other deaf individuals. They are cut off from others; they are isolated. In our culture, I would much prefer to be blind than to be deaf.

In this book I concentrate on those parts of human experience that are primarily auditory: (1) the perception of an object's location in space by sound alone; (2) the perception of nonspeech sounds, particularly music; and (3) the perception of speech. These experiences seem to be part of every human's experience. All cultures have evolved language, all cultures have evolved some type of musical expression; and even newborn infants judge whether a sound comes from the left or right. It seems reasonable to expect that there should be a match between the physical properties of sound, psychological experience, and the physiological properties of the auditory system.

There are basic similarities in the three phenomena. All are concerned with identification of events, ranging from the identification of an individual by voice over a telephone to the identification of a musical instrument (a clarinet as distinct from an oboe). All are concerned with the perception of patterns, ranging from a sentence or song to a melody or the movement of a fly buzzing around your head. And all are concerned with the perception of rhythm. After all, we dance to sounds; we do not dance to lights.

Auditory events tell us about the surrounding world. What is characteristic about these events is that they are spatial and temporal; the information specifying the event consists of variations in the sound over both short and long time periods. These variations often occur almost instantaneously (the changes in loudness, quality, and timbre of a note when a violin is initially bowed); they may also occur in short time periods (the changes in pitch, duration, and timbre for a syllable or a note) or occur in long time periods (the changes in loudness, duration, order, and rhythm

among elements of a sentence or a musical phrase and the changes in position for a moving object).

What will make the analysis difficult is that what happens in the short time intervals affects what happens in the long time intervals and vice versa. We should not think of independent acoustic units that are butted together. Rather, in combining consonants and vowels to form syllables, the articulation forces the acoustic properties of each to invade the other so that both consonant and vowel come out physically different from what they would be if paired with a different vowel or consonant. We hear the same "d" sound in the syllables [da] and [du], although the acoustic signal for "d" changes dramatically. The lack of correspondence between the sound wave and the percept is greater in longer segments. When we listen to an utterance, we hear a sequence of sounds (roughly letters, termed *phonemes*) in which one begins when the previous one ends; each one appears to be conveyed by a discrete packet of sound. Again, this is not the case: there are no units in the acoustic pressure wave clearly separated by physical breaks that correspond to each discrete perceptual unit. The spelling of the word cat is "c" followed by "a" followed by "t." If, however, we try to cut out the "c" part from a tape recording, no unique section can be found. The "c" permeates the entire word acoustically, albeit not perceptually. Similarly, the rhythm of a sentence does not merely come from the accent on each syllable but depends simultaneously on the individual syllables, on the individual words, and on the meaning of the sentence as well. For example, even if each word in the phrase "the white house" were equal as measured by an electronic meter, a listener might report hearing "THE white house" if the utterance specified one of the possible white houses, or "the WHITE house" if the utterance specified the President's lodging, or "the white HOUSE" if the utterance specified which white object. The context as a whole, not each word separately, determines the rhythm and meaning. Unfortunately, it *is* as complicated as it sounds.

The relationship between the physical stimulus and the phenomenal perception is not clear-cut. The phenomenal world of the acoustic events of a listener is not necessarily that described by the physical properties of the sound energy. There is no sound pressure–variation that will always lead to one and only one perception. Similarly, there is no perception that always comes from one and only one pressure variation. If the converse were true—if for every different sound percept there were a unique pattern of sound pressure and if each different sound pressure pattern led to a unique percept—then the problem of auditory perception would be solved, and not by psychologists. It would be solved by physicists who could accurately measure the sound pattern. Perceiving would become rote memorizing; all that would be necessary would be associating each sound pattern with its name or meaning.

This is not the case. Listening is not the same as hearing. The physical pressure wave enables perception but does not force it. Listening is active; it allows age, experience, expectation, and expertise to influence perception. It is often helpful to illustrate how the ear is like a microphone or how the eye is like a camera. It is a mistake, however, to equate the ear with listening or the eye with looking, or to equate the faithful recording of sound energy or light energy with hearing or seeing. We hear and see things and events that are important to us as individuals, not sound waves or light rays (e.g., Noble 1983). Nonetheless, we must measure the physical signal so that we can begin to understand the relationships between the signal as measured by a meter and the event the listener hears. The study of listening must take place within the context of the environment in which listening evolved, since it is the product and reflection of that environment. After all, in spite of the complexities, understanding music and speech comes naturally to all of us.

This orientation should be contrasted to the classical view of sensation and perception. The earliest Greek philosophers held that the mind experiences the external world through the senses. The Greek philosophers, attempting to explain the truthfulness of perception, believed that objects give off little replicas of themselves—eidolas—that when conducted to the mind, allow us directly to perceive the object due to their similarity to that object. In this case, the sensation is the perception. Today we know that acoustic and visual energy stimulate nerves and that it is the nerve firings that are transmitted to the brain. Each nerve responds to one type of energy, and the firing of that nerve is presumed to result in one sensory quality, regardless of the way it was stimulated. These qualities are the sensations, the simple conscious experiences like red, salty, or high pitch. The sensations are the bits and pieces of the perceptions, the blobs of color on a French Impressionist painting, the discrete light bulbs in a scoreboard, or the individual whistles in a bird song. These sensations become the basis for perception. However, many objects will yield the same set of sensations; and conversely, the same object will yield differing sets of sensations, depending on context. Perception is the necessary second stage, the process by which these elements are bound into objects and events. During perception, the conception of an external event is constructed. Perception is based in part on experience, and only through that experience is it possible to make sense of the ambiguous, discrete sensations.

Psychologists have long believed in this dichotomy between sensation and perception. It has led to research that has attempted to determine the sensitivity and accuracy of each sense organ. If the ear is only a microphone, then what is the softest sound it can hear? How sensitive is it to changes in loudness or pitch? How much distortion can it detect? The important point is that this experimentation used the simplest situations,

with sounds that were simple and unchanging. It precluded the discovery of the possibilities for perceiving events and objects. On the whole, the results of this research, ably summarized by Yost and Nielsen (1985), Green (1976), and Moore (1982), will not be presented here.

What I will do is emphasize the perception of events. This first requires a detailed analysis of the physical characteristics of sounds themselves. Most natural sounds are not constant; they change more or less continuously from start to finish, with all parts interacting with one another. Only after the acoustic input is described can we ask questions about the relationships between the physical input and the perceived events. The relationships found between the physical and psychological worlds will motivate and suggest how to look at the physiological world. Here we will ask what sensory and physiological mechanisms exist to explain the psychological, phenomenal experiences. If, for example, we find that many sounds are perceived and identified because of a rapid increase in intensity, then we should look for physiological mechanisms that "fire" to increases in sound pressure or loudness. This organization is consistent with the view that the perceptual systems evolved to cope with the possibilities of environmental stimulation.

This short chapter sets the foundation for the study of listening. Auditory events are set in time and they are perceived in time. For this reason the changes in the characteristics of sound from onset to decay must be explicated. If an understanding of our experiences is possible, it must be correlated to the temporal characteristics of sounds. Before this is possible, however, the basic physical principles underlying sound production must be understood, and this is the topic of the next chapter.

Further Reading

I General Perceptual Theory

Allport, F. H. (1955). *Theories of perception and the concept of structure.* New York: Wiley. Though slightly out of date and eccentric, it is the book I always refer to.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century-Crofts. This book, along with Boring's *A history of experimental psychology* (1950), are the standards for historical background.

Rock, I. (1975). *An introduction to perception.* New York: Macmillan. A good introduction; the focus, however, is visual perception.

II Specific Perceptual Theories

A Ecological Theory The ecological approach to perception developed by Gibson is the dominant theme throughout this book. Michaels and Carello provide an introduction, and Gibson's two books are the "bibles."

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.
Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.
Michaels, C. F., and Carello, C. (1981). *Direct perception.* Englewood Cliffs, N.J.: Prentice-Hall.

B Information Theory

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts.* New York: Wiley. A rewarding discussion of structure.

C Gestalt Theory

Koffka's *Principles* is the textbook of gestalt psychology and Kohler's *Task* provides a recent summary by one of the founders.

Koffka, K. (1935). *Principles of gestalt psychology.* New York: Harcourt, Brace, & World.
Kohler, W. (1969). *The task of gestalt psychology.* Princeton: Princeton University Press.

D Phenomenology

Ihde, D. (1976). *Listening and voice: A phenomenology of sound.* Athens: Ohio University Press.

E Comparisons between Audition and Vision

Geldard, F. A. (1970). Vision, audition, and beyond. In W. D. Neff (ed.), *Contributions to sensory physiology.* Vol. 4 (pp. 1–17). New York: Academic.
Julesz, B., and Hirsh, I. J. (1972). Visual and auditory perception: An essay of comparison. In E. E. David and P. Denes (eds.), *Human communication: A unified view* (pp. 283–340). New York: McGraw-Hill.
Marks, L. E. (1978). *The unity of the senses.* New York: Academic.