Auditory perception

Christopher J. Plack

- **The nature of sound**
  Basic properties.

  *How sound is produced and
    propagated*
  Sound pressure, intensity, and level
  *Pure and complex tones, frequency,
    and the spectrum*

- **From air to ear**
  **How sound is converted into neural
  activity.**

  *Anatomy of the peripheral auditory
    system*
  Frequency analysis in the cochlea
  The role of hair cells—transduction

- **From ear to cortex**
  **How sound is represented and
  analysed by auditory neurons.**

  The auditory nerve
  Rate-place coding
  Phase locking
  *The auditory pathways: ascending
    and descending*

- **Loudness and intensity
  perception**
  Internal representation of intensity.

  Audiograms
  Explanation of loudness
  Detecting changes in intensity

- **Pitch**
  Perception of periodicity.

  *Pure tones: rate-place and
    temporal codes*
  *Complex tones, resolved and
    unresolved harmonics*
  Pattern recognition theory
  Temporal theory
  Modern models and current research

- **Spatial hearing**
  Locating sounds in space.

  Interaural time differences
  Interaural level differences
  Role of the pinna

- **Organisation of sounds**
  **Segregation of the sounds from
  different sources.**

  Simultaneous and sequential grouping
  Continuity illusion
  Sound identification

- **Overview**

  A short summary

For most of us, hearing is central in our interactions with people and with our environment. Despite the rise of the internet, email, and text messaging, speech remains the most important form of human communication, allowing us to communicate our thoughts and feelings with each other with great subtlety. Music is also an integral part of most of our daily lives, whether we listen to it, play it, or are plied with it by advertisers. Furthermore, sounds of all kinds help us interact with the environment. Some sounds warn us of danger, some wake us up in the morning, and others provide information about the operation of machines such as car engines and microwave ovens.

This section will begin with a description of the basic properties of sound, introducing the slightly tricky concept of the *spectrum* that is essential to an understanding of the auditory system. We will then consider how the ear analyses and then encodes acoustic information in the form of nerve impulses that can be processed by the brain. Finally, we will consider our perception of some of the basic acoustic properties that enable us to segregate and identify sounds.

THE NATURE OF SOUND

Production and propagation

Sound is carried by pressure variations in some material, which can be a solid, a liquid, or a gas. Because we are most familiar with sound as pressure variations in the *air*, this section will focus on how sound waves are produced in, and transmitted through, the air.

Molecules in the air are in constant motion. Any material that is in contact with the air experiences a pressure produced by bombardment from the molecules. When an object like a musical instrument or the human vocal cords vibrates, this creates a disturbance in the air. When the object is in an outward phase of vibration, the air molecules are compressed together, and this creates a region of *high pressure*. In the inward phase, the air expands to fill the vacated space, creating a region of *low pressure*. These pressure variations are called *sound waves*.

Sound waves travel through the air around the sound source in a similar way as water waves travel from a disturbance on the surface of a pond. When you drop a pebble into a pond, a ring of waves travels outwards from the initial splash, expanding in two dimensions. Similarly, a disturbance in the air causes sound waves to travel outwards, although in this case the waves radiate in three dimensions around the disturbance. Someone listening to the sound experiences a sequence of pressure variations as the sound waves pass by, just as someone on the edge of the pond would experience a sequence of water waves, with alternating peaks and troughs. At atmospheric pressure, sound waves travel at about 330 metres per second.

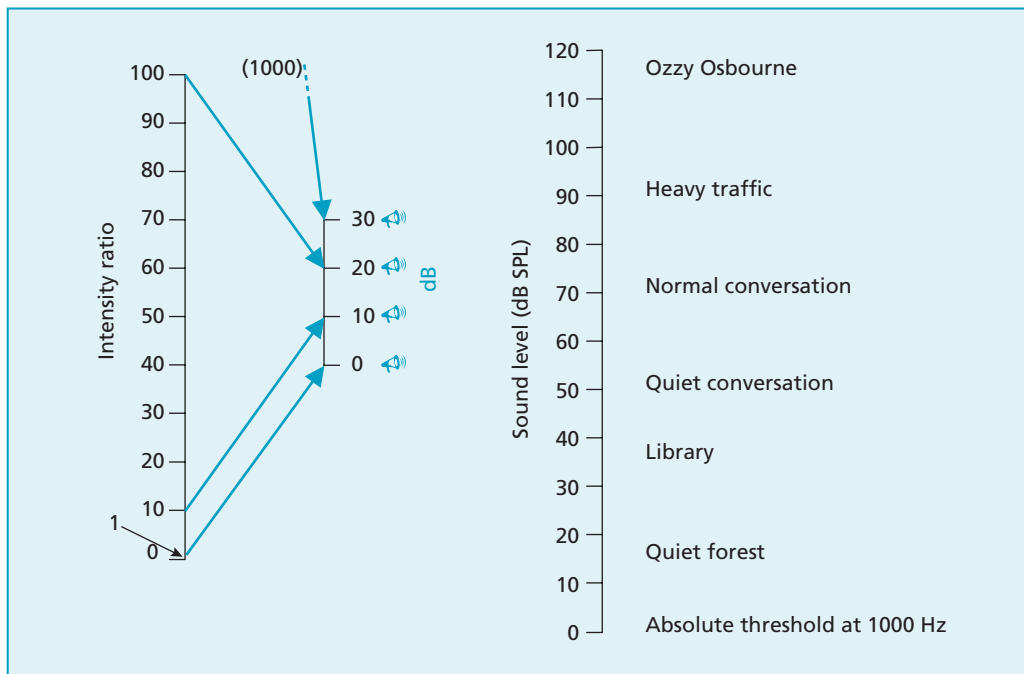### Pressure, intensity, and level

The pressure variations that we usually experience are incredibly tiny. For example, a sound near **absolute threshold** (the lowest sound level that we can hear) has pressure variations less than one *billionth* of atmospheric pressure. This is equivalent in scale to a wave 1 millimetre high on an ocean 1000 kilometres deep. The loudest sound we can listen to without pain has pressure variations a million times greater than this, although these variations are still small compared to atmospheric pressure. The *intensity* of a sound wave (the amount of energy transmitted per second through a given area, for example, a square metre of air) is proportional to the *square* of the pressure. A sound near pain threshold has a pressure a million times greater than a sound near absolute threshold, so the auditory system can operate over a range of intensities of a million million (a million squared)!

Since the range of sound intensities that we hear is so large, sound levels are usually expressed in logarithmic units called **decibels** (dB). A constant increase in dB corresponds to a constant *multiplication* of the sound intensity. An increase by 10 dB corresponds to an increase in intensity by a factor of 10. Similarly, an increase by 20 dB corresponds to an increase in intensity by a factor of 100, and an increase by 30 dB corresponds to an

An illustration of how intensity ratios correspond to the dB scale (left) and an illustration of the sound pressure levels for some familiar environments (right). Click on the loudspeaker symbols to hear sound. © Christopher J. Plack.

increase in intensity by a factor of 1000. The factor of 1,000,000,000,000 in terms of intensity can now be expressed by a range of 120 dB. A scale like this needs a reference point, and 0 dB SPL (sound pressure level) is defined as the level of a sound wave with a pressure of 0.00002 Newtons per square metre, which is roughly the lowest sound level that we can hear.

Pure tones, frequency, and the spectrum

The simplest sound wave from an acoustician's viewpoint is called the *sine wave* or **pure tone**. A pure tone is roughly the sound you get when you whistle a note. A pure tone is characterised by *sinusoidal* variations in pressure with time (a sinusoid is defined by the sine function in mathematics, which cycles back and forth between peaks and troughs). If you whistle from a low note to a high note, the time between the peaks in the sound wave decreases, and the number of peaks in pressure passing a point in space every second increases. The time between peaks is called the *period* of the wave, and the number of repetitions of the **waveform** every second is called the **frequency** of the wave (expressed in cycles per second, or *Hertz*; abbreviated to Hz). The frequency (in Hz) is equal to one divided by the period (in seconds).
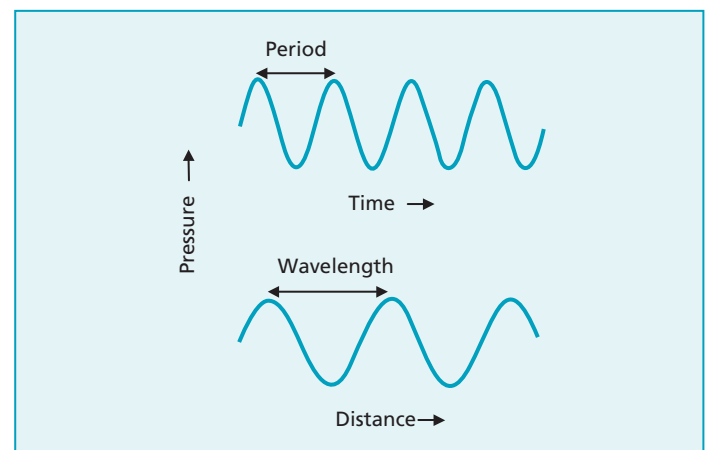
Obviously, most sound waves with which we are familiar are much more complex than simple pure tones. However, any sound wave, no matter how complex, can be produced by *adding together* pure tones with different frequencies. It follows, then, that any sound wave can be described in terms of a number of pure-tone components with different frequencies, levels, and temporal alignments (*phases*). "Bright" sounds, such as the crash of a cymbal, contain mainly high-frequency components. "Warm" sounds, such as a bass drum, contain mainly low-frequency components. (As a rule of thumb, the more rapidly a complex waveform wiggles up and down, the more

The sinusoidal pressure variations over time and distance for a pure tone. © Christopher J. Plack.

intense are the high-frequency components.) The **spectrum** of a sound can be represented by plotting the levels of the different frequency components that are present in the sound. Light too contains different frequency components, which we experience as different colours (red, yellow, green, blue, etc.), and the spectrum of the light reaching us from the sun is dramatically revealed in the rainbow.

A sound wave that repeats over time, such as a vowel sound or the sound made by a tonal musical instrument, has a discrete set of frequency components, each an *integer multiple* of the *repetition rate* of the waveform (the number of times the waveform repeats itself or cycles in a second, also called the **fundamental frequency**). These components are called **harmonics**, and a periodic sound wave that has more than one component is called a **complex tone**. If a complex tone has a fundamental frequency of 100 Hz, for example, the spectrum may contain harmonics with frequencies of 100 Hz (first harmonic), 200 Hz (second harmonic), 300 Hz (third harmonic), 400 Hz (fourth harmonic), and so on. However, it is not necessary that all these harmonics are present for the sound to have the same fundamental frequency. For example, the fundamental frequency is unchanged (although the pattern of the waveform that is repeating is altered) if the first harmonic is removed.
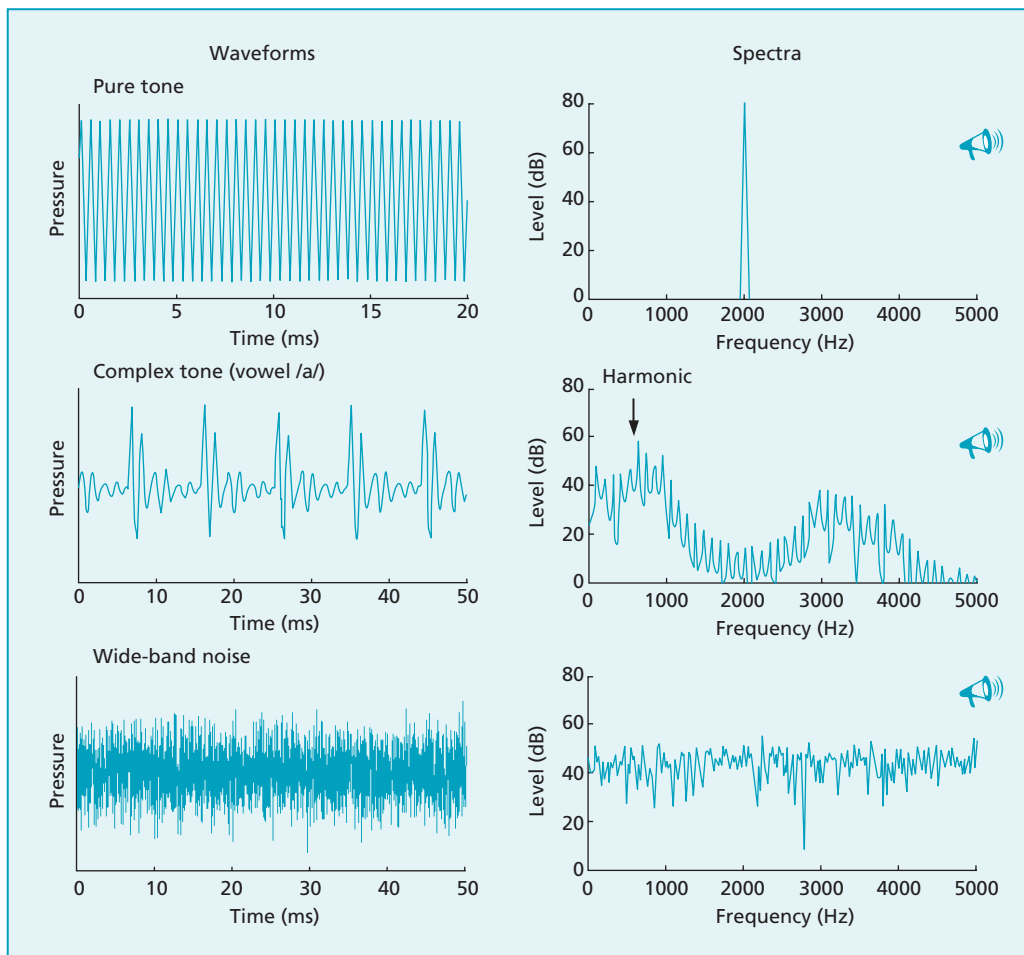

A complex tone can be produced by adding together pure tones at harmonic frequencies. Notice that even if the first harmonic is removed (lower panel) the repetition rate of the complex remains the same (200 Hz, one period every 5 ms). Click on the loudspeaker symbols to hear sound. © Copyright Christopher J. Plack.
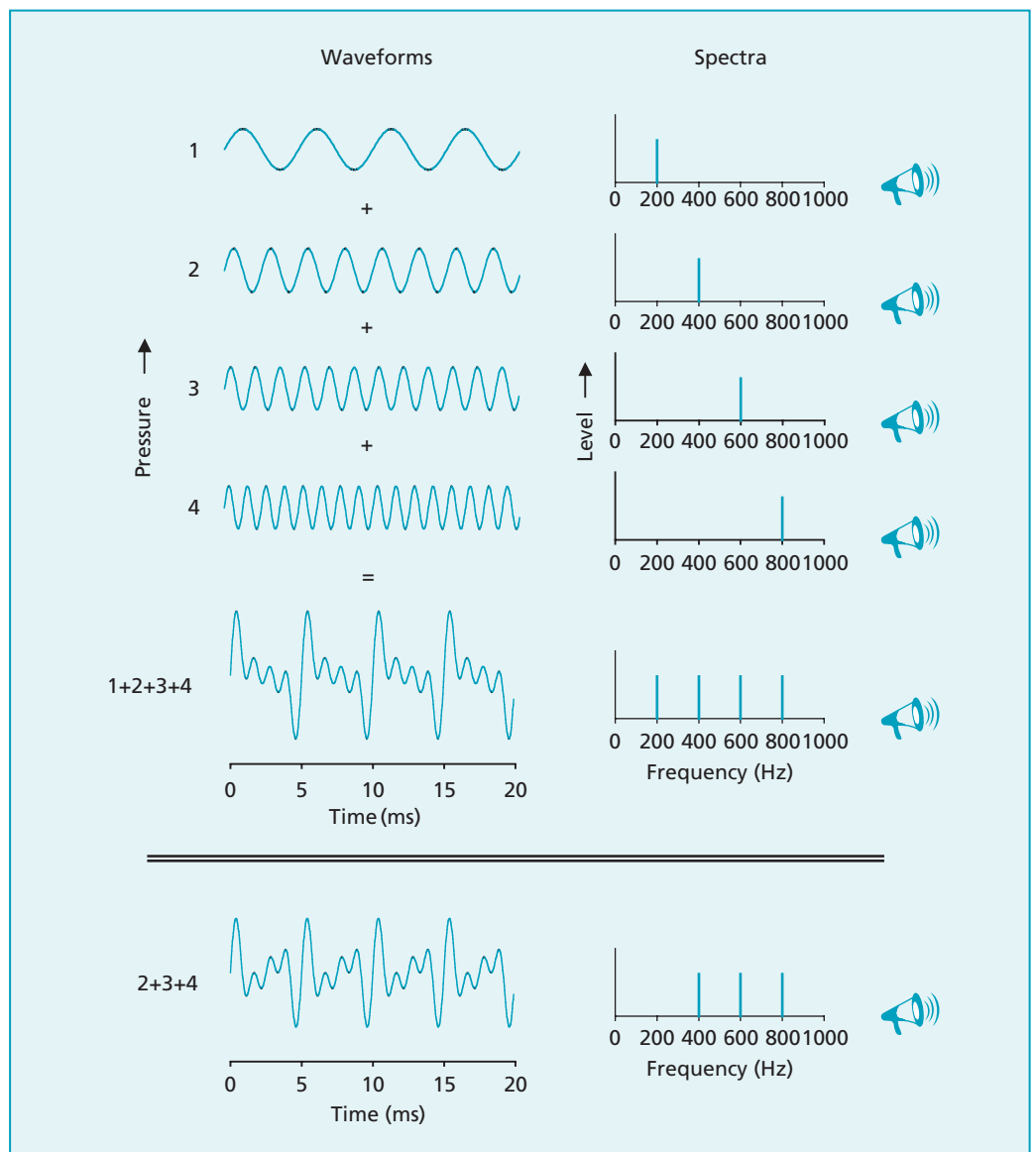
KEY TERMS

Spectrum: the distribution across frequency of the pure-tone components that make up a sound wave.

Fundamental frequency: the repetition rate of the waveform of a complex tone.

Harmonics: the pure-tone frequency components that make up a complex tone.

Complex tone: a tone composed of a number of different pure-tone frequency components, usually with frequencies that are integer multiples of the fundamental frequency.

The waveforms and spectra of three sounds. Notice that the spectrum of the complex tone contains a number of regular peaks corresponding to harmonic frequency components. Click on the loudspeaker symbols to hear sound. © Christopher J. Plack.

A sound wave that does not repeat over time contains a continuous distribution of frequency components. These include impulsive sounds, such as clicks, and noise, which is characterised by a random sequence of pressure variations over time. The background "hiss" of a radio is a good example of noise, as is the sound produced by a waterfall. Noise can be *filtered* to contain a range of frequency components in any part of the spectrum.

FROM AIR TO EAR

The peripheral auditory system

When most people talk about the "ear" they are actually referring to the *pinna*, the cartilaginous flap that protrudes from the side of the head. In fact, the pinnae are not very important to human hearing, although they do help us localise sounds to some extent by modifying the incoming sound waves in different ways depending on the direction of the sound source (see the next page). If someone were to cut your pinnae off, you would *not* be made deaf.

Sound waves enter the ear via the ear canal, a short crooked tube that ends at the *eardrum*. The ear canal has a resonance, just like an organ pipe, and this makes us most sensitive to sound frequencies between about 1000 and 6000 Hz. The eardrum is a taut membrane, and when sound waves reach the eardrum they cause it to vibrate. The more intense the sound, the greater is the amplitude of the vibration. Quiet sounds produce very

small movements: the eardrum moves less than the width of a hydrogen atom for sounds near absolute threshold.

The vibrations of the eardrum are then carried through the middle ear by three tiny bones (the smallest in the body) called the *ossicles* (*malleus, incus*, and *stapes*). The stapes is connected to another membrane: the oval window of the *cochlea*. The cochlea is a narrow, fluid-filled tube, coiled up into a spiral. The cochlea is where the pressure variations are converted, or *transduced*, into electrical activity in neurons of the auditory nerve. It is here that the acoustic information is translated into the language of the brain.

Running along the length of the cochlea is the *basilar membrane*. When the acoustic vibrations arrive at the cochlea they cause pressure variations in the cochlear fluids. These pressure variations cause the basilar membrane to vibrate up and down, though in a quite remarkable way. In an ingenious series of experiments that resulted in the award of the 1961 Nobel Prize in Physiology and Medicine, von Békésy (1947) used stroboscopic light to view the motion of the basilar membrane in ears extracted from dead humans and
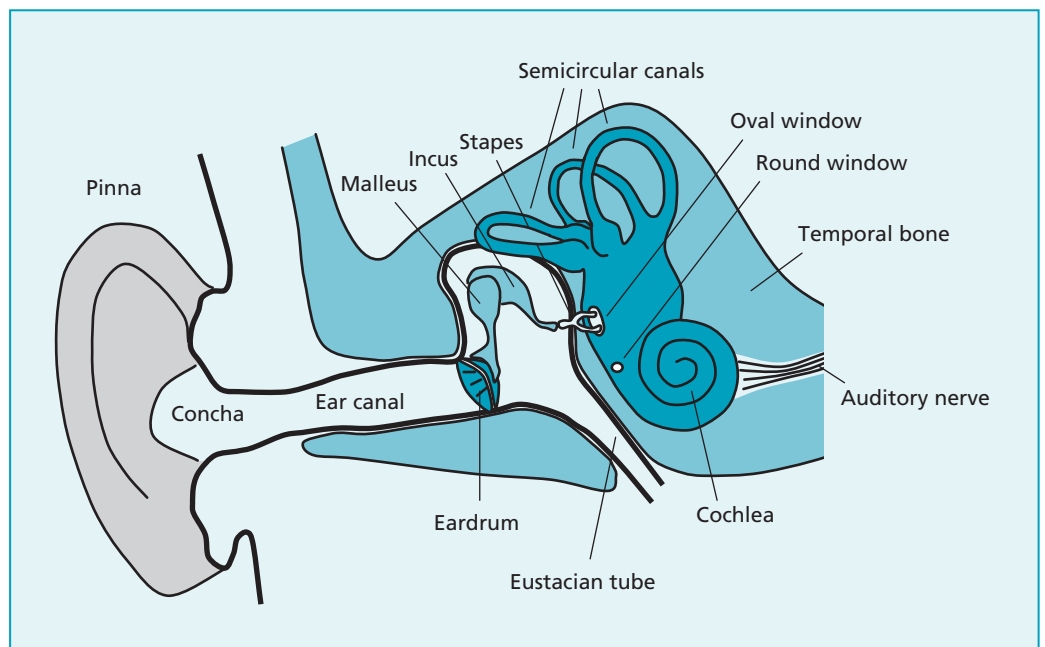
The anatomy of the human ear. © Christopher J. Plack.
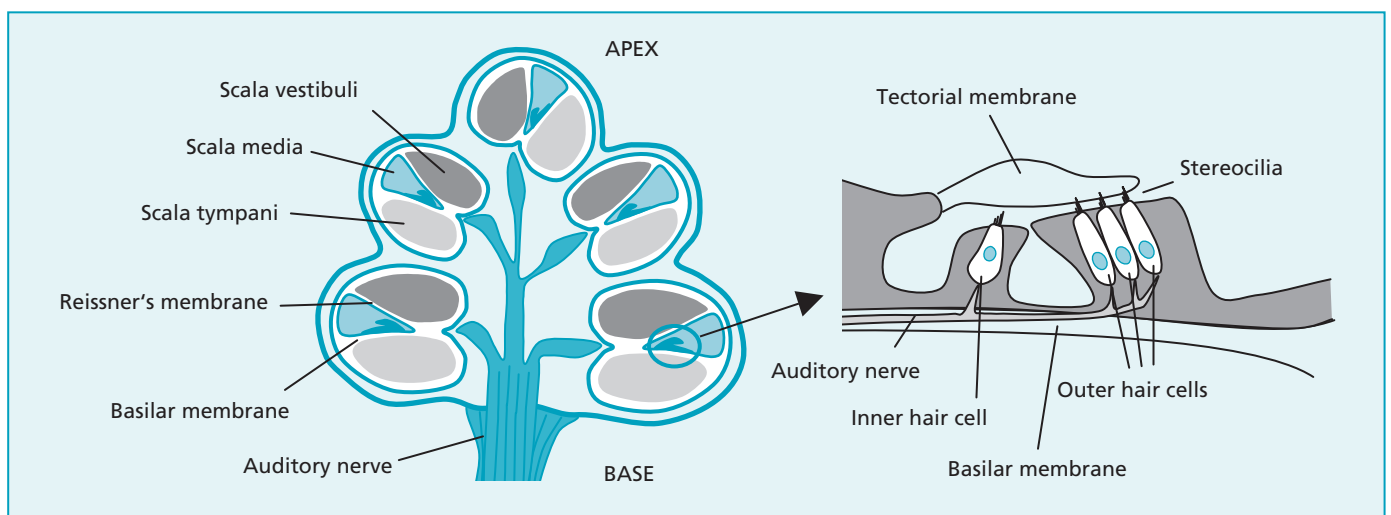
A cross-section of the cochlea. © Christopher J. Plack.

animals. He sprinkled silver particles onto the membrane to improve visibility. When he played a pure tone to these isolated ears it caused a maximum of vibration at a particular place on the basilar membrane, and the place stimulated depended on the *frequency* of the tone. In other words, each place on the basilar membrane is *tuned*, just as each wire on a piano is tuned to play a particular note.

The basilar membrane at the base of the cochlea (closest to the stapes) is thin and stiff and is most sensitive to high-frequency pure-tone components. The basilar membrane near the tip or apex of the cochlea is thick and loose and is most sensitive to low-frequency components. These properties vary continuously along the length of the membrane so that each place is most sensitive to a particular **characteristic frequency**. A given place on the basilar membrane will vibrate most strongly to a pure tone close to its characteristic frequency. As the frequency of the pure tone is moved away from the characteristic frequency the response diminishes. Thus each place on the basilar membrane acts as an **auditory filter**, selectively responding to a narrow range of frequencies. The basilar membrane as a whole can be regarded as a *bank of auditory filters*, each tuned to a different characteristic frequency. The range of frequencies to which each auditory filter responds varies with frequency. Very roughly, each filter will give a high response to a range of frequencies covering 10% of the characteristic frequency.

Frequency selectivity

What do these properties of the basilar membrane mean for the way we hear sounds? Earlier we learned that any complex sound wave can be described in terms of a number of pure-tone components, which, when added together, produce the complex waveform. The main function of the basilar membrane is to separate out (to a limited extent) the different frequency components in a complex sound. When we hear a sound with high-frequency components, such as a cymbal, the basilar membrane vibrates most strongly near the base. When we hear a sound with low-frequency components, such as a bass drum, the basilar membrane vibrates most strongly near the apex. However, the basilar membrane has much more resolution than this example suggests. In fact, the basilar membrane can separate out the *individual harmonics* of a complex tone (for low harmonic numbers, see below).

In a sense, the ear transforms the sound waves entering the ear into a sort of auditory rainbow, just as water droplets (or a prism) separate out the different frequency components in light to give us the visual version. This is a *fundamental* step in the perception of sound. Sounds are identified largely by their spectra, and by the way their spectra change over time. Different speech sounds, such as vowels, have characteristic spectra. Because the basilar membrane represents sounds spectrally, the auditory system is able to detect the differences in frequency composition that identify sounds. Furthermore, our ability to separate out sounds that occur together (for example, two people speaking at once) is largely dependent on our ability to separate out the different frequency components of these sounds. As long as the separation in frequency is sufficient, we are able to "hear out" one frequency component in the presence of another.

As an example, consider two pure tones, one more intense than the other. When the pure tones have a similar frequency, it is hard to hear the quieter tone 🔊 (click on the speaker to hear sound). The quiet tone that you are trying to hear (called the *signal*) is *masked* by the more intense tone (called the *masker*). When the tones have very different
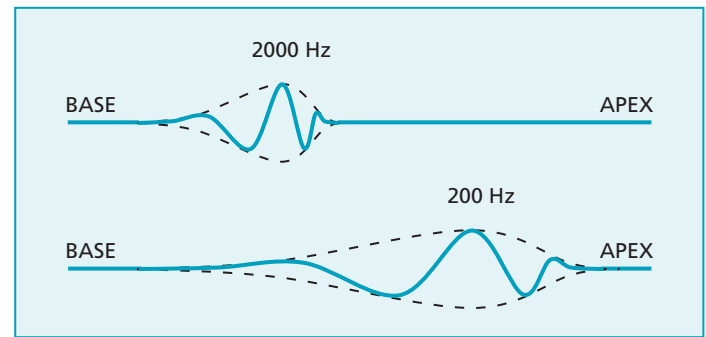
The pattern of vibration on the basilar membrane (stretched straight for clarity) in response to pure tones of two different frequencies. Each place on the basilar membrane that is excited moves up and down, but not in the same phase, so that the resulting pattern appears a little like a wave on water (travelling from base to apex). The dotted lines show the maximum amplitude of the vibration. © Christopher J. Plack.
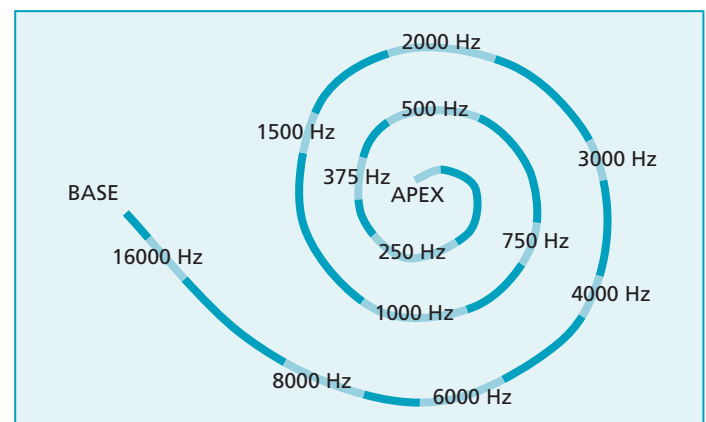
The distribution of characteristic frequencies along the basilar membrane (looking down on the cochlear spiral). © Christopher J. Plack.

How (and why) does the ear break sound down into different frequency components?

KEY TERMS

Characteristic frequency: the pure-tone frequency to which a given place on the basilar membrane, or a given neuron in the auditory system, is most sensitive.

Auditory filter: a term used to refer to the tuning properties (or frequency selectivity) of a single frequency channel in the auditory system. The peripheral auditory system behaves as an array of auditory filters, each passing a narrow range of frequency components.

Helmholtz: The ear as a frequency analyser

Hermann von Helmholtz (1821–1894) was one of the greatest scientists of the nineteenth century. He made important contributions to general physics (particularly the principle of energy conservation), fluid dynamics, electrodynamics, optics (including the invention of the ophthalmoscope), visual physiology, physical acoustics, and the relation between music and auditory physiology. Previously, Joseph Fourier (1768–1830) had proved that it is possible to produce *any* periodic waveform (such as the sound produced by a vibrating string) by adding together pure-tone components with frequencies that are integer multiples of the repetition rate. Helmholtz knew that it was possible to "hear out" these harmonics individually. He proposed that the human ear behaves as a frequency analyser, that can *separate out* the different pure-tone frequency components of a complex sound.

Later experiments have confirmed his theory, and have measured more accurately the limitations of frequency selectivity in humans. Fletcher (1940) conducted masking experiments with noise maskers and pure-tone signals to show that only noise components close to the signal frequency contributed to masking. He coined the term "critical bandwidth" to refer to this range of frequency components. Following Helmholtz, Fletcher suggested that the ear behaves like a bank of overlapping *bandpass filters* (now called *auditory filters*), each tuned to a particular frequency and allowing through a limited range of frequencies. Measurements of basilar-membrane vibration in other mammals correspond well with modern behavioural measures in humans. The results suggest that the frequency selectivity of the human ear is dependent on the mechanical properties of the basilar membrane.

What is psychoacoustics?

The physiological basis of hearing can be studied in other mammals, such as the guinea pig or the chinchilla. Following strict ethical guidelines, physiologists can record from auditory neurons and can make measurements of the response of the basilar membrane. These experiments are limited, however, because they do not inform about the *sensations* that the animal experiences when it listens to sounds. What we really want to know is how our (human) perceptions are related to the physical characteristics of sounds and to the auditory mechanisms that process these sounds. This is where psychoacoustics comes in. Psychoacoustics is the behavioural study of hearing, in which conscious (usually human) listeners are required to make perceptual decisions regarding the sounds that are played to them. This may involve discriminations between sounds that differ in some way (for example, to find the smallest detectable change in frequency, or the lowest level of a signal that can be heard) or it may involve subjective judgements of the loudness or location of a sound, for example. Although psychoacousticians usually find themselves in psychology departments, psychoacoustic techniques can be used to study everything from the motion of the basilar membrane, to neural coding, to more traditional areas of psychology such as memory and attention.

It is possible that the term "psychoacoustics" was first used during the Second World War. A secret US government project was set up to investigate, in part, the potential of acoustic weapons. During the course of the programme, a Dr T.W. Forbes remarked that their field could be called "psycho-acoustics" and the name has stuck (Burris-Meyer & Mallory, 1960). It should be noted that the principal aim of the project was unsuccessful: the acoustic death-beam was not invented!

KEY TERM

Frequency selectivity: the ability of the auditory system to separate out the frequency components of a complex sound.

frequencies, however, it is easy to hear the signal ◀)) (click on the speaker to hear sound). This is because the place on the basilar membrane that responds best to the signal does not respond very well to the masker, as it has a different frequency. It is thought that the auditory system can selectively "listen" to the vibration at a single place on the basilar membrane. In this way we can detect the signal and ignore the masker.

In *psychoacoustic* experiments, listeners are asked to make decisions about sounds that are presented to them, usually over headphones in a soundproof booth. Experiments like these can be used to measure the **frequency selectivity** of the human auditory system. In one technique, a brief (say 4 millisecond) pure-tone signal is fixed at a low level, and played just after a longer duration pure-tone masker. The level of the masker is adjusted until the signal can no longer be detected. When the masker and the signal have the same frequency, the level of the masker needed is low. When the masker is changed to a different frequency, the level of the masker needed is high. This is because the place on the basilar membrane that responds best to the signal responds less well to a masker with a different frequency. Hence, the masker level needs to be higher to compensate. A plot of the masker level needed against the frequency of the masker describes a "tuning curve", which is a measure of the frequency selectivity of the auditory system. Tuning curves like this can be measured for a number of different signal frequencies, thereby characterising the properties of different places on the human basilar membrane.

The hair cells and transduction

Sitting on top of the basilar membrane, and arranged along its length, are rows of *hair cells*. Hair cells have tiny hairs, or *stereocilia*, protruding from their tops. The rows of hair cells closest to the outside of the cochlea are called *outer hair cells*. There are around 12,000 of these cells in the human cochlea (Møller, 2000). Outer hair cells are thought to modify the vibration of the basilar membrane, making us more sensitive to low-level sounds and "sharpening up" the tuning of the membrane so that we are better at separating frequency components. Dysfunction of, or damage to, the outer hair cells is the main cause of deafness.

The innermost row of hair cells consists of the *inner hair cells*, of which there are about 3500 in the human cochlea. These cells convert the motion of the basilar membrane into electrical signals. As the basilar membrane moves up and down, the stereocilia sway from side to side. When they sway in one direction (towards the outside of the cochlea), protein filaments connecting the stereocilia stretch, and this opens tiny plugholes in the *ion channels* in the stereocilia. Positively charged potassium ions enter the cell and cause a voltage change in the cell. The voltage change triggers the release of a chemical neurotransmitter, which diffuses across the tiny gap between the hair cell and a neuron in the auditory nerve. The arrival of the neurotransmitter produces electrical impulses in the neuron. In this way the acoustic signal is converted into an electrical signal that travels up the

auditory nerve to the brain. Inner hair cells are thus the interface between the acoustic world and the perceptual world.

Because each inner hair cell responds to vibration at a single place on the basilar membrane, each inner hair cell is also tuned, responding best to a pure tone at its characteristic frequency. Since each neuron in the auditory nerve is connected to a single inner hair cell, each neuron is also tuned, and carries information about the vibration of a single place on the basilar membrane.

FROM EAR TO CORTEX

The representation of frequency and intensity in the auditory nerve

There are around 30,000 neurons in the auditory nerve, arranged in a bundle. Those neurons in the centre of the bundle originate in the apex of the cochlea and consequently are tuned to low frequencies. Those neurons near the edge of the bundle originate in the base of the cochlea and consequently are tuned to high frequencies. In this way the spatial mapping of frequency along the cochlea is transformed into a spatial mapping of frequency in the auditory nerve.

If the intensity of a sound is increased, the amplitude of vibration on the basilar membrane increases, as does the displacement of the stereocilia on the hair cells. More ion channels are opened, more ions enter the hair cell, more neurotransmitter is released, and the neuron attached to the hair cell produces more impulses per second. In short, an increase in sound intensity produces a *greater rate of firing* in the neurons in the auditory nerve.

Neurons in the auditory nerve cannot fire at rates greater than about 200 impulses per second. Most neurons reach this rate of firing below a level of 60 dB SPL. Above this level they can no longer change their firing rates in response to changes in intensity, and the neurons are said to be *saturated*. However, a few neurons that are a little less sensitive can continue increasing their firing rates up to levels as high as 100 dB SPL. Virtually the entire level range of human hearing can be represented in terms of neural firing rates.

Each neuron in the auditory nerve carries information about a small portion of the spectrum of the incoming sound. The level of these frequency components is represented in terms of the firing rate of the neuron. It follows that the complete spectrum of the sound is carried up to the brain in the auditory nerve, mapped on to the spatial array of auditory neurons. If the sound has more intense high-frequency than low-frequency components, then neurons near the outside of the bundle will fire more strongly than neurons near the centre of the bundle. This is called a *rate-place code*, since the spectral information is represented in terms of firing rate as a function of place in the neural array.
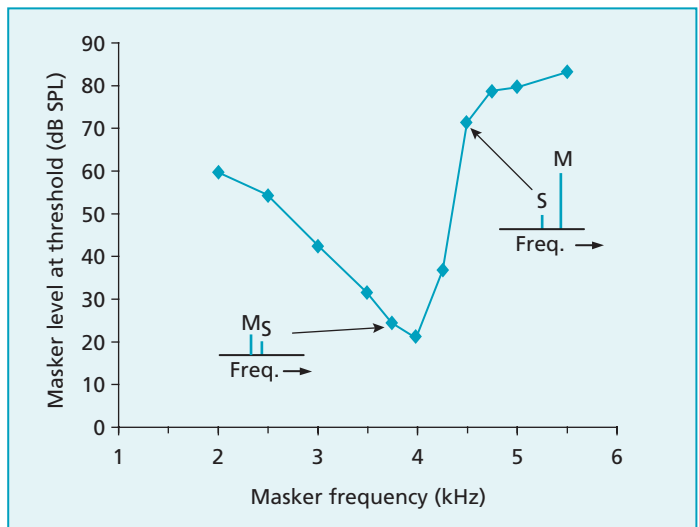
A tuning curve, showing the masker level needed to mask a 4000 Hz pure tone signal (S) as a function of masker (M) frequency. The greater the frequency separation between the masker and the signal, the less masker excitation on the basilar membrane at the place tuned to the signal, hence the higher the masker level required. Data from Yasin and Plack (2003). © Christopher J. Plack.

How is sound intensity represented in the auditory system?

Cochlear hearing loss and hearing aids

The hair cells in the cochlea are very sensitive and are susceptible to damage by intense sounds and disease. Drugs such as kanamycin (an antibiotic) and carboplatin (a chemotherapy drug) can also damage the hair cells (Møller, 2000). In high doses, even aspirin can cause transient hearing loss. Furthermore, as we get older there is a gradual deterioration in hair cell function called *presbycusis*, which mainly affects sensitivity to high frequencies. Cochlear hearing loss is the most common type of deafness in the UK, affecting about one in five people. Although dysfunction of the inner hair cells disrupts the transduction process, making us less sensitive to the motion of the basilar membrane, the main cause of cochlear hearing loss is dysfunction of the *outer hair cells*.

Listeners with hearing loss have increased absolute thresholds (inability to hear low-level sounds), and testing for this is the usual means of diagnosis. However, listeners with *cochlear* hearing loss also have poor frequency selectivity and an abnormally rapid growth in loudness with level. Healthy outer hair cells are necessary for normal frequency selectivity and for the amplification of low-level sounds that leads to a shallow growth in basilar membrane vibration with sound level. The abnormal growth of loudness means that although a hearing-impaired listener may not be able to hear low-level sounds, high-level sounds seem just as loud to them as they do to normal-hearing listeners. A hearing aid that simply amplifies all sound levels will allow the user to hear low-level sounds but will make high-level sounds unbearably loud! Many modern hearing aids employ some sort of *compression* or *gain control*, so that they amplify more at low levels than at high levels (like the outer hair cells). Unfortunately, hearing aids cannot restore the reduction in frequency selectivity. Listeners with cochlear hearing loss find it hard to separate out the different sound components entering their ears. This makes it hard to listen to a person speaking in a noisy room, for example. Since there is at present no remedy for this aspect of hearing loss, the message is simple: avoid loud sounds and look after your ears!
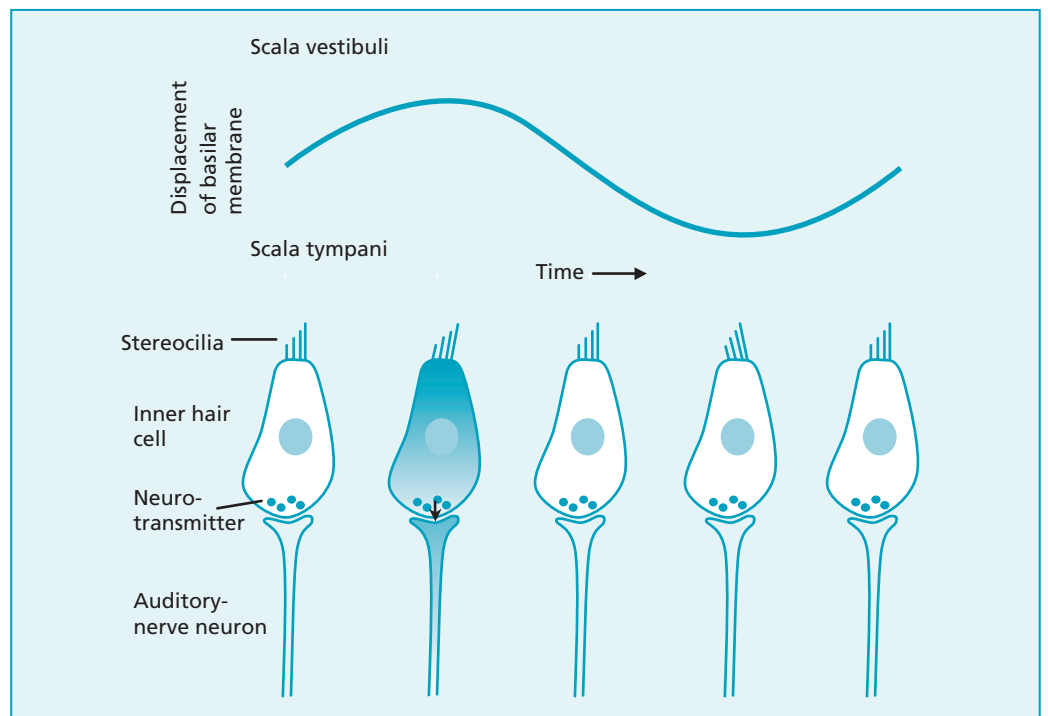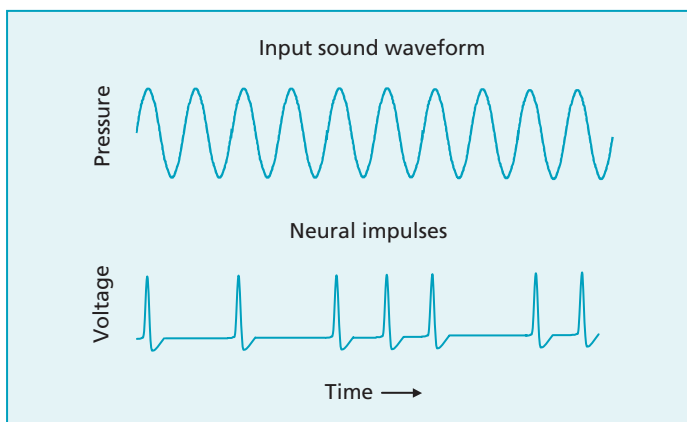
Electrical activity in the inner hair cells and in the auditory nerve is related to the motion of the basilar membrane. Stimulation occurs when the stereocilia are bent towards the outside of the cochlea (to the right in this illustration). © Christopher J. Plack.

Phase locking

There is another way in which information about a sound is carried in the auditory nerve. Earlier we learned that an inner hair cell is stimulated when its stereocilia are bent in one direction, away from the centre of the cochlea. What this means is that more neurotransmitter is released, and neural impulses are more likely to be produced, when the basilar membrane is moving in one direction. Since sound waves involve alternating high pressure and low pressure, the basilar membrane moves up and down in response. However, nerve impulses are more likely to be produced when the basilar membrane moves just one way. It follows that the pattern of firing is *phase locked* to the vibration of the basilar membrane, and hence to the pressure variations in the air.

Phase locking in an auditory-nerve neuron in response to a pure tone. Notice that although the neuron is not firing on every cycle of the tone, impulses are synchronised with a particular phase of the waveform. © Christopher J. Plack.

Take the example of the pure tone again. When a 100 Hz pure tone is played to the ear, the basilar membrane will vibrate up and down at 100 Hz, with the maximum vibration near the apex of the cochlea. Neurons connected to this part of the membrane will tend to produce impulses that are separated by 10 ms (or by integer multiples of 10 ms), which is the period of a sound wave with a frequency of 100 Hz. Although neurons cannot fire at rates much greater than 200 impulses per second, they will continue to phase lock at much higher frequencies. At these higher frequencies neurons will not fire on each and every cycle, but the impulses will still tend to occur periodically, at the same time in the cycle of the waveform. It is thought that this temporal synchrony occurs in response to pure tones with frequencies up to 5000 Hz. Above this frequency, neurons cannot phase lock to the individual variations in pressure. However, they will phase lock to peaks in the overall amplitude, or *envelope*, of the sound wave. In this case, the neurons are more likely to fire when the sound wave has a higher overall amplitude.

Phase locking implies that each neuron in the auditory nerve not only carries information about the gross level of activity at the place in the cochlea from which it originates, but also carries information about the precise pattern of vibration of the basilar membrane at that place. As we will discover later, this information is used for pitch perception and for locating sounds in space.

The ascending auditory pathway

The auditory nerve travels from the cochlea to the brainstem. The brainstem nuclei are groups of neurons, mirrored in each hemisphere, that perform analysis on the incoming information. The first processing station is the *cochlear nucleus*. The neural signal is then passed to the superior olivary complex, which also receives input from the opposite, or *contralateral*, ear. It is thought that the superior olivary complex is involved in sound localisation. Further up the pathway are the *lateral lemniscus* and the *inferior colliculus*. Each stage contains many different types of neurons with varying properties that decode and analyse the auditory information. The system is massively interconnected, with parallel pathways carrying different types of information. The information is passed to the primary auditory cortex via the *medial geniculate body* of the thalamus. The primary auditory cortex is located at the top of the temporal lobe, mostly hidden within the *Sylvian fissure*.

The organisation of frequency in terms of place is carried on all the way up the pathway to the auditory cortex. Just as different places in the visual cortex respond to light from different points in space, different places on the auditory cortex respond to
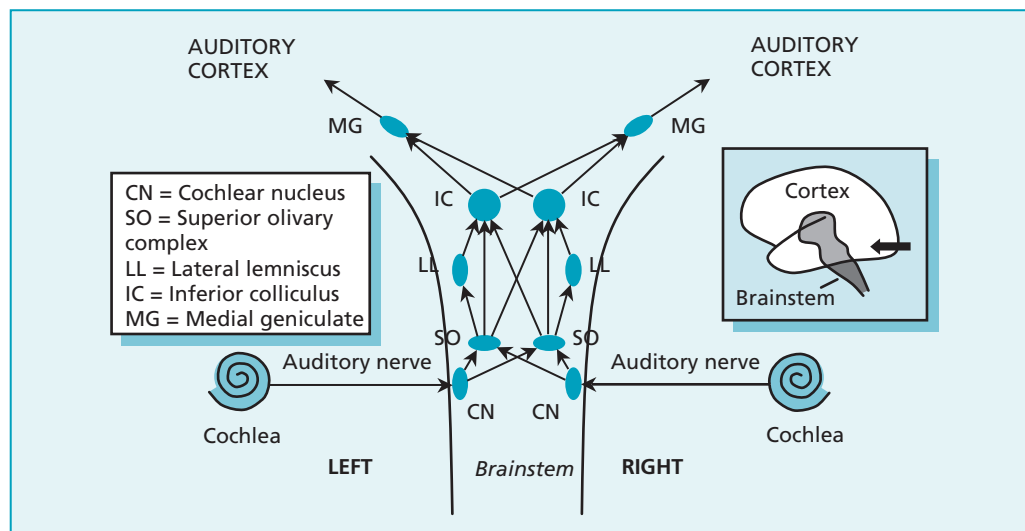
The main connections in the ascending auditory pathway. The viewpoint is towards the back of the brain, as indicated by the arrow on the illustration to the right (cerebellum removed). © Christopher J. Plack.
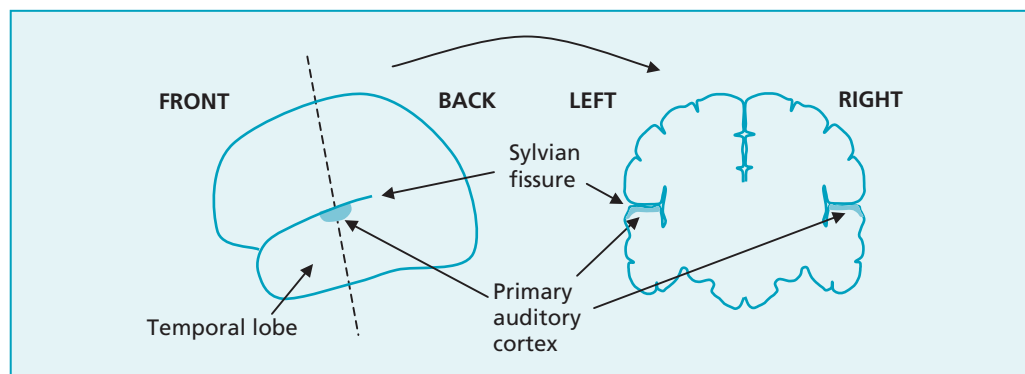
The location of the primary auditory cortex, shown from the side and in a cross-section taken along the dotted line. © Christopher J. Plack.

different frequencies. However, some neurons have complex tuning properties with peaks of sensitivity at multiple frequencies, and some have evidence of lateral inhibition, so that certain frequencies cause the firing rate of the neuron to *decrease*. These neurons have different tuning properties to neurons in the auditory nerve. It is clear that they receive input from neurons with different characteristic frequencies, and combine the signals in various ways. Neurons such as these may be involved in processing spectral information across frequency and across time. Understanding exactly how this processing relates to our perceptions is of great interest at present.

In addition to the ascending pathway, there is also a *descending* auditory pathway, carrying signals from the cortex, through the brainstem nuclei, to the outer hair cells in the cochlea. It is possible that the cortex may use this pathway, not only to influence the processing in the brainstem, but also to influence the motion of the basilar membrane. When the *efferent* neurons that innervate the outer hair cells are stimulated, the basilar membrane becomes less sensitive to low sound levels and frequency selectivity is reduced (Guinan, 1996). The exact purpose of this mechanism is yet to be determined, but it is an intriguing aspect of mammalian hearing.
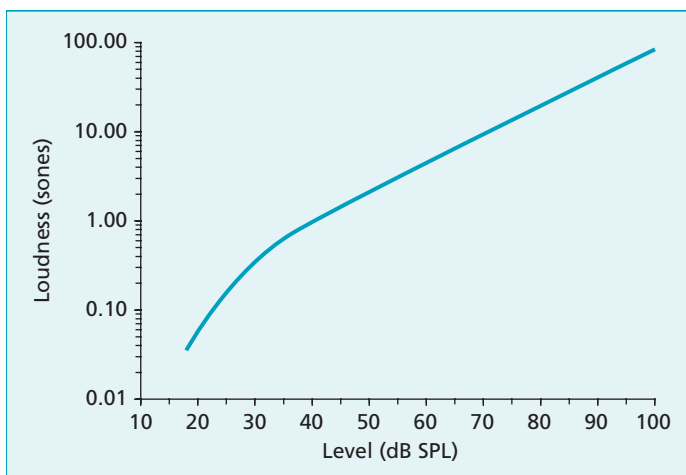
LOUDNESS AND INTENSITY PERCEPTION

As we have seen, the human auditory system is very sensitive, being able to detect tiny variations in pressure that produce tiny displacements of the eardrum. It is possible to plot the absolute threshold of a pure tone as a function of frequency. This is called an *audiogram*, and is used by audiologists to test for hearing loss. We are most sensitive to frequencies in the range 1000–6000 Hz, and sensitivity falls off at low and high frequencies (the tone needs to be more intense for us to detect it at these extremes). The highest frequency that can be heard by humans with excellent hearing (generally, young children) is about 20,000 Hz, although even for these listeners the sound level needs to be very high. As we get older we tend to lose our sensitivity to high frequencies.

Loudness refers to the subjective magnitude of a sound. The word "loudness" should not be used to refer to a physical property of a sound. A sound may have a *physical level* of 80 dB SPL, and we *experience* this as a particular loudness. Loudness is the perceptual correlate of sound level: as the level increases so does our experience of the magnitude of the sound. However, the relation is not straightforward. Using techniques in which listeners are required to compare the loudness of two sounds (for example, to adjust the level of sound A until it seems to be twice as loud as sound B), it has been shown that loudness scales with intensity to the power of about 0.2–0.3. Put another way, to double the perceived magnitude of a sound, you have to increase its level by about 10 dB, which is a factor of 10 in intensity. For the electric guitarists among you, if you have a 100 watt amplifier and you want one that sounds twice as loud, then you have to upgrade to a 1000 watt amplifier!

The shallow function relating loudness to intensity implies that the huge range of intensities to which we are exposed is compressed into a much smaller range of subjective magnitudes. This compression is mainly the result of the outer hair cells, which effectively amplify the vibration of the basilar membrane in response to low-level sounds, but not high-level sounds. This means that the large range of sound intensities is mapped on to a much smaller range in terms of the intensity of vibration on the basilar membrane.

We have seen how the intensity of sounds is represented in terms of the firing rates of auditory neurons, and this is largely the information that gives rise to the sensation of loudness. It remains a bit of a mystery, however, as to why we are as good at detecting small changes in

The relation between loudness and sound level. Loudness is expressed in units called *sones*, where one sone is defined as the perceived magnitude of a 1000 Hz pure tone presented at a level of 40 dB SPL. If a sound appears twice as loud as this reference tone, then it has a loudness of two sones. Based on Hellman (1976). © Christopher J. Plack.

intensity at high levels as we are at low levels. Up to levels of at least 110 dB SPL, when most auditory-nerve neurons are not contributing useful information because they are saturated, we can detect changes in the level of a wideband noise (a noise with a wide range of frequency components) of just 0.5 dB (Miller, 1947). Therefore, under experimental conditions, we can just detect the difference between a sound with a level of 100 dB SPL and a sound with a level of 100.5 dB SPL. One possibility is that we do not use all the level information in the auditory nerve. The relatively few neurons that contribute at high levels may be sufficient for intensity coding. The greater number of neurons that contribute at low levels may provide more information than we actually need, and at some stage in the auditory pathway this information may be discarded (see Plack & Carlyon, 1995, p. 142).

It has been shown that we are good at detecting changes in the spectral shape of sounds. For example, we can detect a "bump" in the spectrum of just a few dB. Indeed, we can discriminate a flat spectrum from a spectrum with a small bump, even when the over-all levels of the two sounds are randomised. The ability to make intensity comparisons across frequency is called *profile analysis* (Green, 1988), and it is vital for identifying sounds. Speech sounds such as vowels are characterised by the locations of spectral peaks and dips.

Finally, the auditory system is good at following rapid changes in the level of a sound over time. We can detect fluctuations in the level of a sound up to rates of 1000 times a second (Bacon & Viemeister, 1985). This high temporal resolution contrasts with the comparatively low temporal resolution of the visual system, for which the equivalent maximum detectable rate of flicker is only around 50 cycles per second. The visual system is "sluggish" compared to the auditory system. We need the high temporal resolution in hearing because many sounds are characterised by rapid changes in acoustic features. This is particularly true of speech. In the visual system, on the other hand, objects are mainly characterised by differences in their *spatial* properties: a high degree of *temporal* resolution is not necessary.

PITCH

A sound wave that repeats over time is often heard as having a distinct **pitch**, which corresponds to the repetition rate. Like loudness, pitch is a perceptual variable, and should not be used to refer to a physical property of a sound (although it often is). Variations in pitch are associated with musical melodies, and with the *intonation contour* of the human voice. Pitch has fascinated hearing researchers for hundreds of years, but despite all the interest, the nature of the auditory mechanisms that underlie our perceptions is still a matter for intense debate.

What is the difference between frequency and pitch?

Pure tones

Pure tones are the simplest repetitive sound waves. For these sounds, the repetition rate is equal to the frequency of the tone. Between about 30 Hz and about 5000 Hz, we can use variations in the frequency of pure tones to carry musical melodies. As we have seen, there are two ways in which the frequency of a pure tone is represented in the auditory system:

1. *Rate-place code*. Frequency is coded by the place on the basilar membrane (and in other regions of the auditory pathway) that is activated.
2. *Temporal code*. Frequency is coded by the pattern of phase-locked firing in the auditory nerve.

The upper limit of phase locking is around 5000 Hz, which is also around the upper limit on our ability to perceive musical melodies (Attneave & Olson, 1971). If a familiar melody is played using pure tones that are above 5000 Hz, the melody is not recognisable. You can hear that something is varying, but it doesn't sound musical. This suggests that there may be a relation between phase locking and pitch. Furthermore, our ability to detect a frequency difference between two pure tones is much better than would be

> **KEY TERM**
> **Pitch:** that aspect of sensation whose variation is associated with musical melodies.

predicted on the basis of the rate-place code. We can detect a frequency difference of just 0.2% for frequencies below 4000 Hz (Moore, 1973). The auditory filters are too broad for this level of resolution. The frequency discrimination threshold—the smallest detectable difference in frequency between two pure tones—increases rapidly at higher frequencies, again suggesting that without phase locking we cannot determine pure-tone frequency very accurately. It appears, therefore, that the pitch of pure tones is derived from the pattern of phase locking, although it is possible that the rate-place representation contributes in some way, so the two systems may be working together (Loeb, White, & Merzenich, 1983).

What would happen to your hearing if your neurons lost the ability to phase lock?

Complex tones

Complex tones also repeat over time, but the spectrum of a complex tone contains a number of harmonics. Because the width of the auditory filters in the basilar membrane is, roughly, a constant proportion of their characteristic frequencies, the filters become broader as frequency increases. At high frequencies, components need to be further
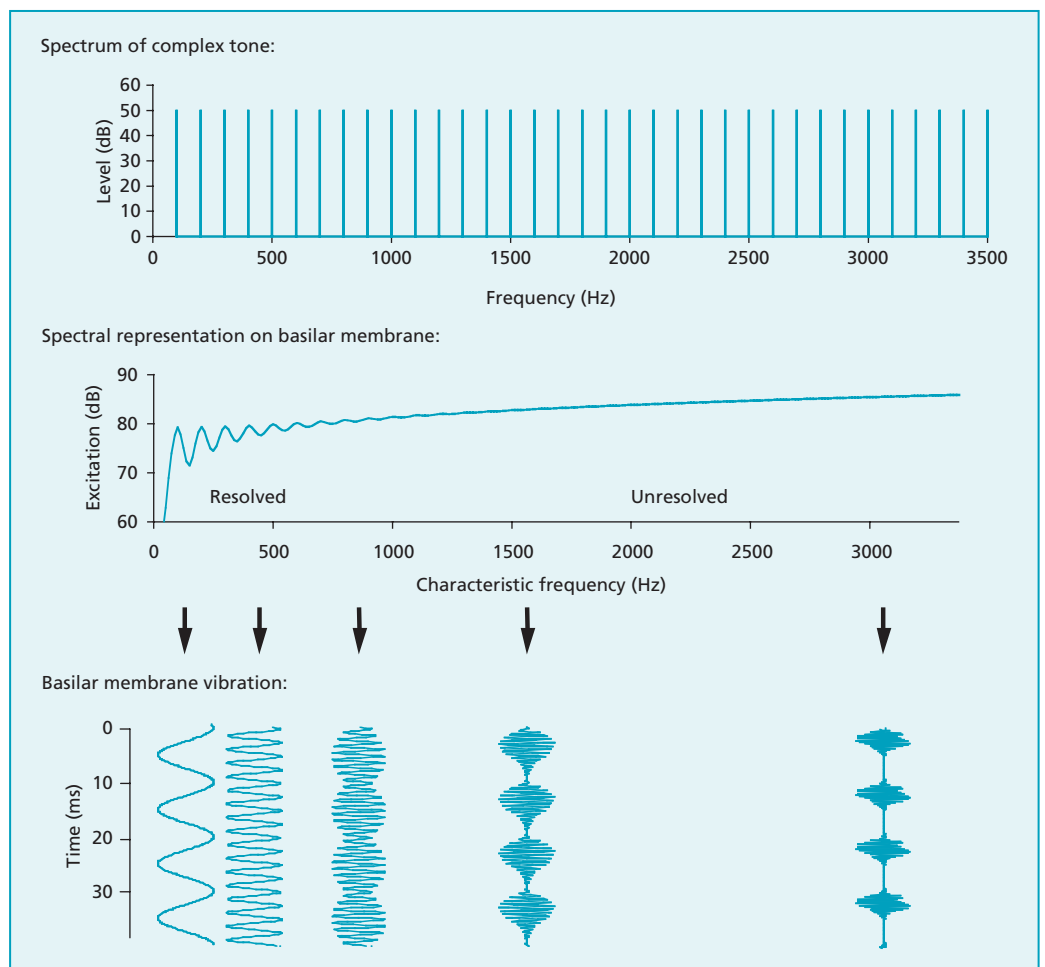
The spectrum of a complex tone consists of a number of pure-tone frequency components (harmonics) that are regularly spaced in frequency (top panel). The middle and lower panels show how this spectrum is represented in the cochlea. The activity of the basilar membrane is illustrated with respect to the characteristic frequency of each place on the membrane. In this plot, the base of the membrane is to the right and the apex to the left. Low-harmonic numbers are separated out by the basilar membrane and there are distinct peaks of activity. A place on the basilar membrane tuned to a low harmonic shows a sinusoidal vibration corresponding to the waveform of the harmonic (lower panel, left). The higher harmonics are not separated into individual pure tone components. Several harmonics interact at each place on the membrane. The result is a complex pattern of vibration (lower panel, right) that has the same repetition rate or fundamental frequency as the original complex tone. © Christopher J. Plack.

apart in frequency to be separated by the membrane. It follows that only about the first 10 harmonics in a complex tone are separated out or *resolved* by the cochlea, and excite distinct places on the basilar membrane. A place on the membrane tuned to a low harmonic shows a pattern of vibration corresponding to the waveform of the pure-tone harmonic.

In contrast, the higher harmonics are unresolved, and several harmonics interact at each place on the membrane. Because of this, the vibration of the membrane is more complex. Indeed, the pattern of vibration of each place is itself a complex tone. Of course, the frequency spacing of the harmonics that interact at each place is the same as the harmonic spacing for the complex tone as a whole. It follows that the repetition rate or fundamental frequency of the vibration of the membrane is the same as that of the original complex tone.

Classical theories suggested that the pitch of a complex tone may be derived in one of two ways:

1. *Pattern recognition theory* (Goldstein, 1973; Terhardt, 1974). The **resolved harmonics** form a pattern that is characteristic of any fundamental frequency. If harmonics of 300, 400, and 500 Hz are present, the auditory system can deduce that the fundamental frequency is 100 Hz. This mechanism requires that the harmonics are resolved, so that their frequencies can be independently determined.
2. *Temporal theory* (Schouten, 1940, 1970). Pitch may be derived directly from the repetition rate of the waveform produced by the interacting unresolved harmonics.

It turns out that a pitch can be heard when *only* resolved harmonics, or *only* unresolved harmonics are present. This means that neither of the classical theories can be a sufficient explanation. However, the lower resolved harmonics tend to dominate pitch perception. Complex tones containing just unresolved harmonics produce a weak (albeit musical) pitch. A slight change in the frequency of a resolved harmonic has a much larger effect on pitch than a change in the frequency of an unresolved harmonic (Moore, Glasberg, & Peters, 1985). Furthermore, we can be cued to hear a pitch corresponding to the fundamental frequency when just two successive resolved harmonics are presented to opposite ears (Houtsma & Goldstein, 1972). Each cochlea only receives one harmonic; therefore interaction between several harmonics on the basilar membrane is clearly not necessary for the low pitch to be heard.

Some modern models suggest that the pitch of complex tones is derived by combining the information from the resolved and unresolved harmonics (Meddis & Hewitt, 1991; Moore, 2003, p. 223). These models can produce a good description of pitch phenomena by assuming that the phase-locking information is analysed across characteristic frequency. Take, for example, a neuron tuned to the third harmonic of a pure tone with a fundamental frequency of 100 Hz. This neuron will phase lock to the harmonic frequency of 300 Hz. If it produces an impulse on every cycle, the impulses will be separated by 3.33 milliseconds (1/300 seconds). However, the interval between every three impulses will be 10 milliseconds, which is the period of the fundamental frequency. In fact, for any resolved harmonic there will be intervals between impulses that correspond to the fundamental frequency. A neuron with a higher characteristic frequency, tuned to unresolved harmonics, will phase lock to the envelope of the waveform produced by the interacting harmonics. The intervals between impulses will also tend to be 10 milliseconds. If the auditory system simply searches across characteristic frequency for the most common interval between impulses, then this is likely to correspond to the period of the complex.
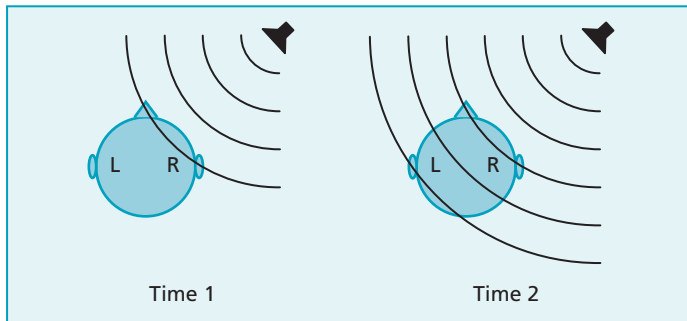
It must be said, however, that the issue is far from settled, and pitch remains an exciting field of research. The pattern recognition models have experienced something of a resurgence recently (Bernstein & Oxenham, 2003), and there has been a suggestion that there may be *separate* pitch mechanisms for the resolved and unresolved harmonics (Carlyon & Shackleton, 1994). If true, then this will seriously complicate our endeavours!

In parallel with the psychophysical investigations, auditory physiologists have been searching for the neural mechanisms that may be responsible for extracting pitch from

periodic waveforms. Much of this research has focused on how the temporal pattern of phase locking is analysed in the brain. The maximum frequency to which a neuron will phase lock is much reduced in the medial geniculate body of the thalamus; neurons will not synchronise to tones with frequencies above about 800 Hz (Møller, 2000). Consequently, physiologists are looking for neurons in the brainstem that are responsible for decoding the phase-locking information into something like a rate-place code. The hope is to find neurons that produce their highest firing rates when they are presented with a characteristic *fundamental* frequency. There is some evidence that there are neurons like this in the inferior colliculus (Langner & Schreiner, 1988).

SPATIAL HEARING

The visual system has remarkable spatial resolution, with millions of receptors in each eye responding to light from different locations in the visual field. In contrast, the auditory system focuses on *frequency* resolution, with thousands of receptors tuned to different frequencies. However, having two ears (effectively, two directional receptors) helps us to locate the direction of a sound source to some extent. There are two main cues available:

1. *Interaural time differences*. A sound to the right arrives at the right ear before the left ear.
2. *Interaural level differences*. A sound to the right is more intense in the right ear.

A sound from the right arrives at the right ear first and is louder in the right ear. These are the main cues for sound localisation. © Christopher J. Plack.
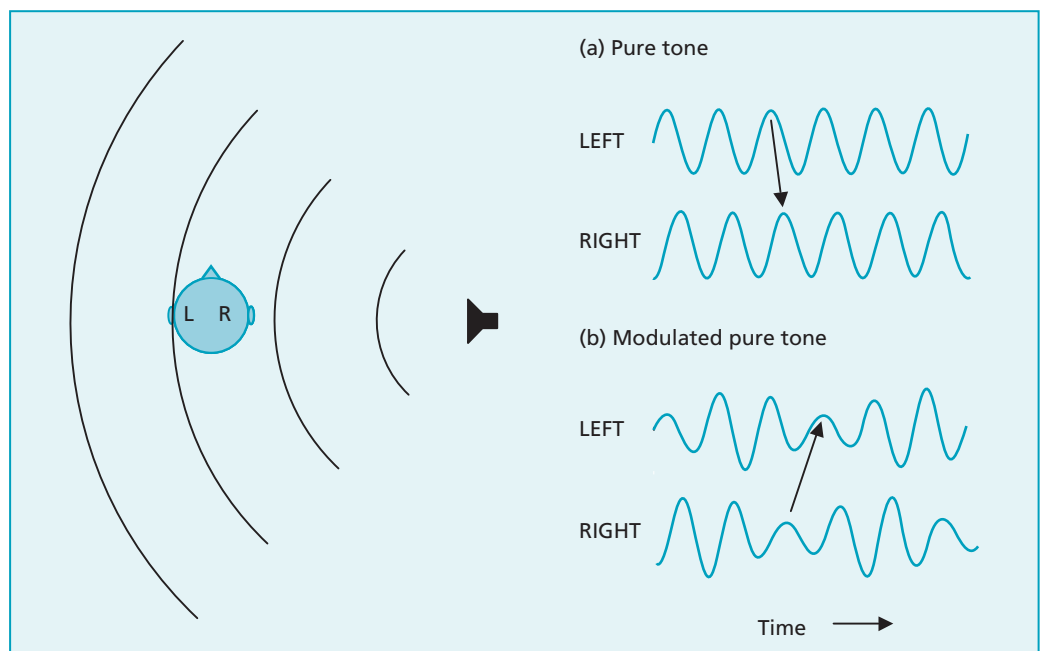
Why do we need two ears?

Interaural time differences appear to be dominant for sound localisation. We can detect an arrival time difference between the two ears of just *10 millionths* of a second (Klump & Eady, 1956)! This corresponds to a direction difference of about 1° relative to straight ahead. If the tone frequency is a little above 750 Hz, the distance between waveform peaks (the *wavelength*) is less than twice the distance between the ears. This means that for a continuous pure tone directly to the right, a waveform peak in the left ear will

Even though the pure tone comes from the right, it appears as if the waveform peaks lead in the left ear. If the tone is modulated in amplitude, the ambiguity is resolved. © Christopher J. Plack.

KEY TERM

Interaural time difference: a difference between the arrival times of a sound wave at the two ears.

be closely followed by a peak in the right ear. In this way it may appear that the sound is to the left when it is really to the right. This ambiguity can be resolved if the waveform envelope fluctuates, so that the more slowly varying features can be compared. In general, however, interaural time differences are more important at *low* frequencies.

Interaural time differences are probably extracted in the medial superior olivary complex. Neurons have been located there that respond to input from both ears, and have a higher rate of firing when there is a particular time delay between the two inputs (Carr & Konishi, 1990; McAlpine & Grothe, 2003). The astonishing resolution of the system implies that the timing of waveform features is very precisely encoded in the phase-locked response of the auditory nerve.

In contrast, interaural level differences are more useful at *high* frequencies. The head will tend to prevent the sound waves from reaching the ear furthest from the sound, creating a sort of "acoustic shadow", and leading to large level differences between the ears. At low frequencies, however, the sound can bend round or *diffract* around the head, so that the level differences are minimal.

Although the time and level cues provide important information about the location of a sound, that information would remain ambiguous were it not for the help of the pinna. For example, a sound directly in front produces the same interaural time and level differences as a sound directly behind, or indeed a sound directly above. The pinna modifies the sound as it enters the ear, and produces peaks and dips in the spectrum that depend on the direction of the sound (Blauert, 1997, p. 67). The auditory system can recognise these "directional signatures" to help localisation. We also use head movements to help resolve ambiguities. If a sound is directly in front and the head is turned to the right, then the sound will arrive at the left ear first. If the sound is directly behind, then the same head movement will cause the sound to arrive at the right ear first.
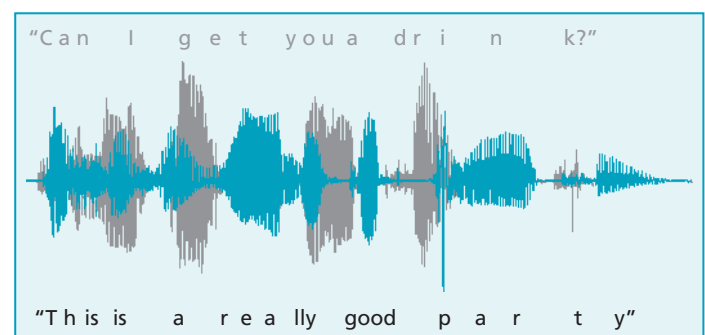
ORGANISATION OF SOUNDS

It is quite rare that the totality of the sound we hear comes from just one sound source. In most cases there are several sound sources, and the sound waves from these sources are simply summed together in the air to produce a highly complex waveform. The auditory system has the difficult task of separating the sounds emanating from the source of

How does the ear separate out sounds that occur together?

interest (for example, the speech from the person you are talking to) from the sounds coming from the other sources in the environment (for example, the other people at the party). Bregman (1990) has termed the whole process **auditory scene analysis**.

There are two aspects to this ability. First, the auditory system must parse the sound components from different sources that occur simultaneously. Second, it must be able to follow the sequence of sounds from the source of interest over time. The frequency selectivity of the cochlea is crucial for the first aspect. The auditory system separates out the different frequency components in the cochlea, and then extracts and groups together those components that come from the source of interest. To decide which components come from which source, the auditory system uses two main rules of thumb:

"Can I get you a drink?"

"This is a really good party"

The waveforms of two sentences spoken at the same time. To follow one of these sentences, the auditory system has to separate out the relevant speech sounds when they overlap with the competing speech, and group together the sequence of sounds over time. © Christopher J. Plack.

1. *Onset times*. Sound components that start together come from the same source.
2. *Harmonicity*. Sound components that are harmonically related (i.e., from a single harmonic series) come from the same source. It is much easier to separate two voices if the fundamental frequencies are different (Culling & Darwin, 1993).

To follow a sequence of sounds from the same source over time, called *sequential streaming*, the auditory system looks for similarity between the sounds. If the sounds are similar in terms of their spectra, their fundamental frequencies, and their locations, then they tend

> **KEY TERM**
> **Auditory scene analysis:** the perceptual organisation of sounds according to the sound sources that are producing them.

to get grouped together perceptually. In this way we can attend to the sequence of sounds from a single sound source over time. Another aspect of this process gives rise to an auditory "illusion" called the continuity illusion. If a sound was interrupted briefly by a second more intense sound (for example, speech interrupted by the sound of a door slamming), then it appears to us as if the first sound was continuous during the interruption. Similarly, if a brief gap in a sound is filled with a more intense sound, then the first sound seems continuous even though it is actually discontinuous. The sensation we experience is an interpretation, or "best guess", made by the auditory system on the basis of the physical evidence. When there are several sound sources competing with each other, much of what we "think" we hear is actually a reconstruction by the auditory system.

Once the sounds of interest have been separated from the acoustic background, they can then be identified. There are many different acoustic features to which the auditory system is sensitive, and these apply to different types of sound. Fundamental frequency is clearly very important for music, as is *timbre*, the sensation relating to the spectral and temporal features that help characterise different instruments. For speech sounds, the spectrum, and the way the spectrum changes over time, provides the main source of information. By focusing on a spectral analysis of sounds, the auditory system can decipher the complex acoustic code that is the basis for human communication.

OVERVIEW

The defining characteristic of mammalian hearing is the spectral analysis performed by the cochlea, which separates out the different frequency components of sounds. This analysis is the basis for much subsequent processing, including pitch perception for complex tones, and sound segregation and identification. The auditory system also encodes the temporal patterns of vibration on the basilar membrane in the form of phase-locked firing patterns in the auditory nerve. The precision of the temporal code gives us accurate pitch information, and enables us to localise sounds on the basis of interaural time differences. The detailed spectro-temporal analysis, and the ability to follow rapid changes in the characteristics of sounds, provides us with a highly efficient and sensitive sensory system.

FURTHER READING

- Moore, B.C.J. (2003). *An introduction to the psychology of hearing* (5th ed.). London: Academic Press. A comprehensive yet readable account of auditory perception, including the basic physiology of the ear.
- Yost, W.A. (2000). *Fundamentals of hearing: An introduction* (4th ed.). New York: Academic Press. Good detail on the physics of sound, as well as anatomy, physiology, psychophysics, and hearing loss.
- Bregman, A.S. (1990). *Auditory scene analysis: The perceptual organisation of sound*. Cambridge, MA: MIT Press. The definitive work on perceptual organisation in hearing.

REFERENCES

Attneave, F., & Olson, R.K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, *84*, 147–166.

Bacon, S.P., & Viemeister, N.F. (1985). Temporal modulation transfer functions in normal-hearing and hearing-impaired subjects. *Audiology*, *24*, 117–134.

Bernstein, J.G., & Oxenham, A.J. (2003). Pitch salience: Harmonic number or harmonic resolvability? *Journal of the Acoustical Society of America*, *113*, 3323–3324.

Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*. Cambridge, MA: MIT Press.

Bregman, A.S. (1990). *Auditory scene analysis: The perceptual organisation of sound*. Cambridge, MA: Bradford Books/MIT Press.

Burris-Meyer, H., & Mallory, V. (1960). Psycho-acoustics, applied and misapplied. *Journal of the Acoustical Society of America*, *32*, 1568–1574.

Carlyon, R.P., & Shackleton, T.M. (1994). Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms? *Journal of the Acoustical Society of America*, *95*, 3541–3554.

Carr, C.E., & Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuroscience*, *10*, 3227–3246.

Culling, J.F., & Darwin, C.J. (1993). Perceptual separation of simultaneous vowels: Within and across-formant grouping by F0. *Journal of the Acoustical Society of America*, *93*, 3454–3467.

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics*, *12*, 47–65.

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.

Green, D.M. (1988). *Profile analysis*. Oxford, UK: Oxford University Press.

Guinan, J.J. (1996). Physiology of olivocochlear efferents. In P. Dallos, A.N. Popper, & R.R. Fay (Eds.), *The cochlea* (pp. 435–502). New York: Springer.

Hellman, R.P. (1976). Growth of loudness at 1000 and 3000 Hz. *Journal of the Acoustical Society of America*, *60*, 672–679.

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

Klump, R.G., & Eady, H.R. (1956). Some measurements of interaural time difference thresholds. *Journal of the Acoustical Society of America*, *28*, 859–860.

Langner, G., & Schreiner, C.E. (1988). Periodicity coding in the inferior colliculus of the cat: I. Neuronal mechanisms. *Journal of Neurophysiology*, *60*, 1799–1822.

Loeb, G.E., White, M.W., & Merzenich, M.M. (1983). Spatial cross correlation: A proposed mechanism for acoustic pitch perception. *Biological Cybernetics*, *47*, 149–163.

McAlpine, D., & Grothe, B. (2003). Sound localization and delay lines—do mammals fit the model? *Trends in Neurosciences*, *26*, 347–350.

Meddis, R., & Hewitt, M. (1991). 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.

Miller, G.A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *Journal of the Acoustical Society of America*, *19*, 609–619.

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. (2003). *An introduction to the psychology of hearing* (5th ed.). London: Academic Press.

Moore, B.C.J., Glasberg, B.R., & Peters, R.W. (1985). Relative dominance of individual partials in determining the pitch of complex tones. *Journal of the Acoustical Society of America*, *77*, 1853–1860.

Møller, A.R. (2000). *Hearing: Its physiology and pathophysiology*. New York: Academic Press.

Plack, C.J., & Carlyon, R.P. (1995). Loudness perception and intensity coding. In B.C.J. Moore (Ed.), *Handbook of perception and cognition: Vol. 6. Hearing* (pp. 123–160). Orlando, FL: Academic Press.

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

Schouten, J.F. (1970). The residue revisited. In R. Plomp & G.F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing* (pp. 41–54). Leiden, The Netherlands: Sijthoff.

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

von Békésy, G. (1947). The variations of phase along the basilar membrane with sinusoidal vibrations. *Journal of the Acoustical Society of America*, *19*, 452–460.

Yasin, I., & Plack, C.J. (2003). The effects of a high-frequency suppressor on tuning curves and derived basilar membrane response functions. *Journal of the Acoustical Society of America*, *114*, 322–332.