12.5. SUGGESTIONS FOR FURTHER READING

General references

J. Frisby (1979) *Seeing*. Oxford University Press. (Provides a well-illustrated introduction to many of the issues discussed in this chapter.)

R. Held, H. W. Leibowitz, H.-L. Teuber (eds.) (1978) *Handbook of Sensory Physiology*. Vol. VIII, *Perception*. Heidelberg: Springer. (A fairly up-to-date compendium on a wide range of topics.)

Specific references

Psychophysical investigation of channels. O. J. Braddick, F. W. Campbell & J. Atkinson (1978) in Held, Leibowitz & Teuber.

Visual function outside cortical area 17. D. C. Van Essen & J. H. R. Maunsell (1983) *Trends in Neuroscience*, **6**, 370.

Texture discrimination and segregation. B. Julesz (1981) *Nature*, **290**, 91.

Pattern recognition. H. B. Barlow, R. Narasimhan & A. Rosenfeld (1972) *Science*, **177**, 569. (A review of some general problems and approaches.)

Scene analysis. P. H. Winston (1977) *Artificial intelligence* New York: Addison-Wesley. (A good view of the programming issues); M. Boden (1977) *Artificial Intelligence and Natural Man* Hassocks; Harvester (Artificial Intelligence in a broader context.)

Constancy of visual direction. Ch. 7 in Howard, I. P. (1982) *Human Visual Orientation*. Chichester: Wiley.

'*Perceptual hypotheses*'; R. L. Gregory (1970) *The Intelligent Eye* London: Weidenfeld & Nicholson.

Selection in perception. D. E. Broadbent (1971) *Decision and Stress* London: Academic Press.

Visual search. P. M. A. Rabbitt (1978) in *Psychology Survey No. 1*, ed. B. M. Foss. London: Allen & Unwin.

CHAPTER THIRTEEN

Basic physics and psychophysics of sound

E. F. EVANS

Though people generally value their sight most highly, hearing is arguably the most important sense for man. For what marks out *Homo sapiens* from other species is his ability to express ideas and concepts, and these he communicates to his fellows chiefly by means of language. This communication occurs first by means of sound, and oral communication remains foremost throughout life. It has been said that a blind person is cut off from the world of *things*, whereas one who is deaf is cut off from the world of *people*.

Yet our understanding of the mechanisms of hearing is less well developed compared with that of the other senses. This has resulted partly because the organ of hearing is so inaccessible – buried, in man, in the *petrous temporal* bone; partly because the important parameters of acoustic signals are less obvious, than, say, those of visual signals, and are measured with considerable technical difficulty and expense; and partly because anaesthetics apparently interfere more with the behaviour of the auditory system than with that of other sensory systems. On the other hand, the application of such basic knowledge as we have on the functioning of the peripheral auditory system has already led to valuable diagnostic tools, and to specific aids designed to compensate for disorders of hearing. There is much more to be known and done.

The multiplicity of terms, measures and scales used in acoustics and psychoacoustics may appear confusing and, at first sight, unnecessary. However, the distinctions are real and important! This first chapter is therefore intended to assist the reader, unfamiliar with the jargon and concepts of acoustics and psychoacoustics, to obtain the necessary background and an understanding of the basic properties of the hearing process.

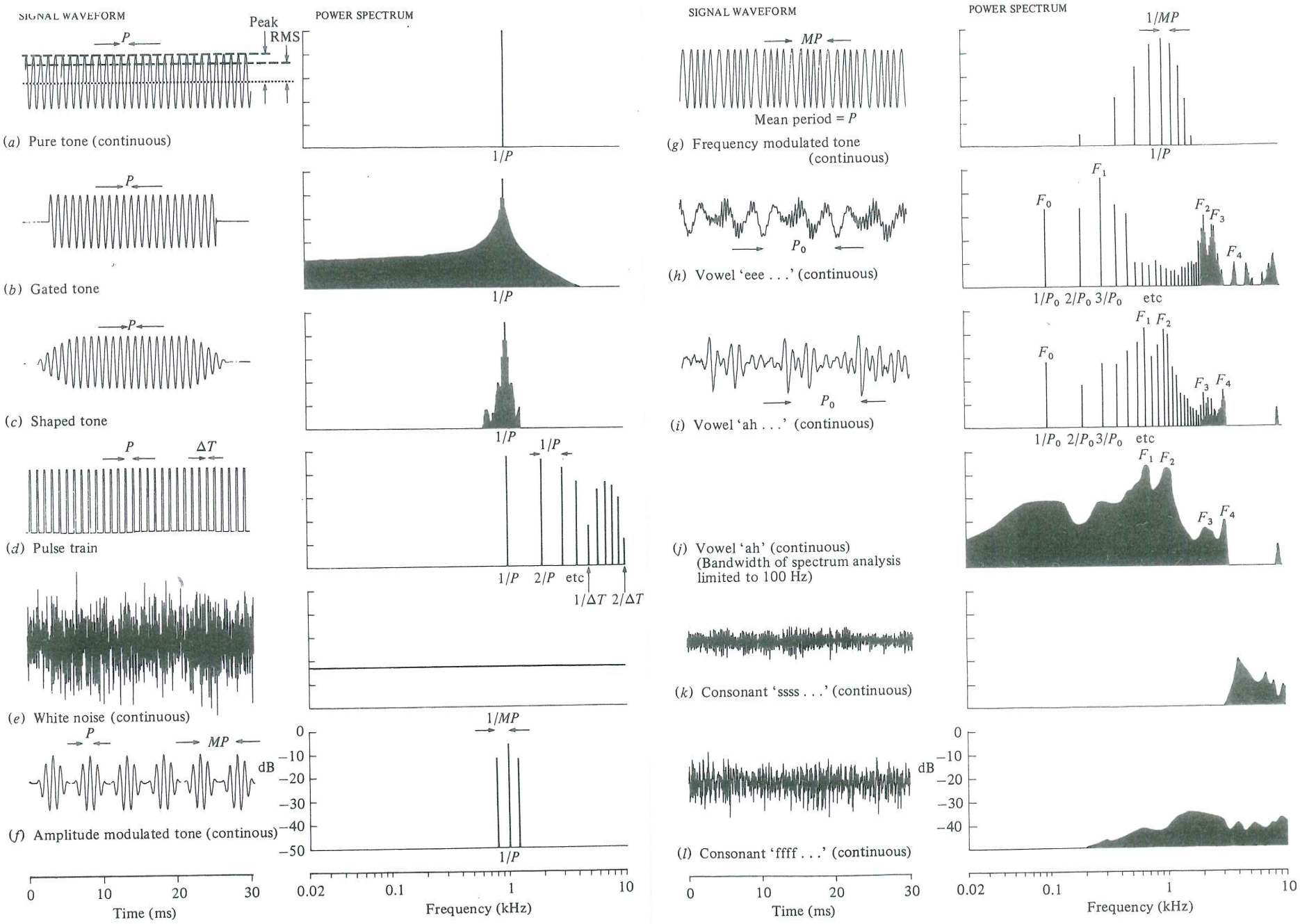


Fig. 13.1. Waveforms and frequency spectra of different signals. The waveforms (on left) describe the excursions of signal (voltage, pressure) with time. Peak indicates the peak amplitude; RMS the root mean square value. *P* indicates the period of the waveform; *MP* the period of modulation. The spectra (on right) describe the results of frequency (Fourier) analysis of each signal, i.e.: the relative level of the individual frequency components contained in the signal, plotted on logarithmic power (dB) and frequency scales.

13.1. BASIC PHYSICAL PARAMETERS OF SOUNDS AND THEIR MEASUREMENT

Any vibration, transmitted to the ear via the air or directly via the bones of the head, is in principle capable of generating auditory sensations. In practice, because of the limitations of the hearing mechanism, and the complex transformations that such vibrations undergo in the ear, only a restricted range of sounds is audible. It is therefore important for us to consider what are the relevant parameters of sound vibrations for our ears, and how they are expressed.

Air-borne 'sounds' are relatively small fluctuations in air pressure. At the lowest intensities we can hear, these fluctuations are incredibly small: about 1×10^{-13} of an atmosphere. The fluctuations may be *periodic* (e.g. Fig. 13.1*a–d, f–i*), i.e. be regular in time, or *aperiodic*, either random (e.g. Fig. 13.1*e*) or impulsive. Periodic sounds generate musical, tone-like percepts; random sounds, 'noise'; and impulsive sounds, click sensations.

For various reasons, it is convenient to measure the magnitude of acoustic signals by their *power*. Furthermore, because of the immense dynamic range of the ear (greater than 10^{12}:1), it is convenient to express these powers on a logarithmic scale. This has the added advantage that some important psychophysical and physiological functions are very roughly proportional to the logarithm of the stimulus power. The log unit of power is named the Bel (after Alexander Graham Bell) and it is conveniently divided into ten *decibels* (abbreviated dB). It is important to remember that decibels are not *units* of measurement in the usual sense, but represent *ratios* of powers. Therefore the decibel difference between the magnitude of two signals is given by $10 \times \log_{10} (I_1/I_2)$, where I_1 and I_2 are their respective powers (intensities), or in terms of their fluctuations in pressure (P_1 and P_2 respectively), $20 \times \log_{10} (P_1/P_2)$.*

Thus, a doubling of pressure equals an increase of level of 6 dB; doubling of power: 3 dB. Usually a signal strength is indicated in decibels *relative to some standard.* This is its *level.* In acoustics, the standard adopted internationally is the sound pressure of 20 μPa (1 Pascal = 1 N m^{-2}), roughly the threshold of hearing for a 3 kHz tone (see Figs. 13.2 and 13.3). Acoustic signal levels are conventionally expressed relative to that standard as *dB SPL* (Sound Pressure Level). Table 13.1 shows, for reference, the relationships between the magnitudes of sound pressure, power (intensity), and SPL, relevant to audition.

* Acoustic power varies as the *square* of pressure. $20 \times \log_{10}(P_1/P_2) = 10 \times \log_{10}(P_1^2/P_2^2)$.

Table 1. *Relations between sound pressure, power, and level, with typical examples drawn from common situations*

Sound pressure (N m^{-2} or Pa)	Power (intensity) (W m^{-2})	Sound pressure level (dB SPL, i.e. referred to 20 μPa)	Examples and some effects (approximate only)
200	100	140	Jet engine; over-amplified rock group; threshold for pain
20	1	120	Damage to cochlear hair cells
6.32	10^{-1}	110	Threshold for discomfort
2	10^{-2}	100	Motor cycle engine
			Orchestra
6.32×10^{-1}	10^{-3}	90	*fff*
2×10^{-1}	10^{-4}	80	*ff*; busy traffic; shouting
2×10^{-2}	10^{-6}	60	*mf*; normal conversation
2×10^{-3}	10^{-8}	40	*pp*; quiet office
6.32×10^{-4}	10^{-9}	30	*ppp*; soft whisper
2×10^{-4}	10^{-10}	20	Country area at night
2×10^{-5}	10^{-12}	0	Threshold of hearing of young person at 1–5 kHz
6.32×10^{-6}	10^{-13}	−10	Threshold of cat's hearing (1–10 kHz)

All temporal waveforms can be analysed mathematically or physically into their component *frequencies.* The breakdown of complex sounds into their constituent frequency parts is known as *frequency analysis.* It can be accomplished mathematically by *Fourier analysis* (or by digital filtering carried out electronically) or physically by means of *filters* (having resonant elements or circuits). The results of such an analysis on some typical waveforms is shown in Fig. 13.1. The concept is important because the ear itself is capable of a limited but vital frequency analysis (see Chapter 1, p. 15).

The limitations of any frequency analysis are the duration of the signal requiring to be analysed and the bandwidth* of the filters used.

* Bandwidth denotes the range of frequencies accepted by a filter without substantial attenuation; see Fig. 15.1*b*.

Any practical frequency analysis is a compromise between frequency and time resolution. At one extreme is Fourier analysis (equivalent to infinitely narrow filtering of infinitely long signals) as in Fig. 13.1*a*, *d*, *e*, *f*, *g*, *h*, *i*). This yields a *long-term spectrum*, where the frequency components are infinitesimally narrow, hence are said to be *line spectra* as shown on the right-hand side of the figure. Towards the other extreme is the *short-term spectrum*, where the time of occurrence of frequency components is more or less preserved at the expense of lack of precision in the determination of frequency. This is achieved by filtering with finite filter bandwidth: the wider the bandwidth, the poorer the frequency resolution, but the better the resolution of signal components in time (and vice versa). The compromise adopted by the ear seems well-adapted to our needs, for the amplitudes of resolvable frequency components change slowly enough to be followed by the neurons of the auditory system. An analogous comparison, in space rather than time, may perhaps explain the advantage of the local spatial frequency analysis thought by some to be performed in the visual cortex (see Chapter 8, p. 150, and also Chapter 1, p. 17).

Sufficiently rapid changes in a tonal stimulus either in intensity or in frequency produce 'splatter' of energy to other frequencies. In the case of switching a pure tone (Fig. 13.1*b*) on and off, the energy is spread over a wide frequency range. This spread can be minimised by appropriate shaping of the onset and offset of the signal (Fig. 13.1*c*).

Brief pulses produce 'clicks' which, like broad-band (white) noise (Fig. 13.1*e*), contain energy across a wide range of frequencies. In the case of repeated clicks (Fig. 13.1*d*), however, the analysis shows the energy as confined to frequencies (*harmonics*) that are multiples of the click repetition frequency (known as the *fundamental frequency*), as in Fig. 13.1*d*. (Note that because the clicks have finite duration, the energy in the harmonics is not uniform, but there exist minima at frequencies corresponding to the reciprocal of the pulse duration, $1/\Delta T$, $2/\Delta T$...etc.)

Periodic changes in a tonal stimulus, either in frequency or intensity, produce 'splatter' of energy to *discrete* neighbouring frequencies as shown in Figs. 13.1*f* and *g* respectively. The spacing of these frequencies is equal to the modulation rate (1/MP), and in the case of frequency modulated tones (Fig. 13.1*g*), the number of the adjoining frequencies increases with the depth of modulation.

Speech sounds (Figs. 13.1*h–l*) are produced by the action of the

many resonances of the vocal tract (mouth, pharynx, nasal passages) on sounds produced by pulsatile excitation of the larynx ('voiced' speech sounds, such as vowels), or by the tongue or lips ('unvoiced' sounds such as most consonants). For continuous steady sounds such as long *vowels* this produces virtually line spectra representing the first 50 or so harmonics of the voice *fundamental frequency* (F_0, Figs. 13.1*h*, *i*). Depending upon the shape of the vocal tract and consequently on the speech sounds being uttered, the amplitudes of the harmonics are emphasised at certain frequencies known as *formants* (F_1, F_2, F_3, F_4 in Figs. 13.1*h–j*). The relative spacing of these formants helps to characterise the particular vowel, and their movements help to distinguish certain consonants. The frequencies spanned by the voice fundamental and formants are indicated in Fig. 13.3.

In the production of (unvoiced) consonants, the energy is relatively wide-band, although again, the distribution of energy is an important cue for the identification of the consonant, as demonstrated in the differences between the spectra of 'sss' and 'fff' in Figs. 13.1*k* and *l* respectively.

As already mentioned, the fineness of the frequency analysis depends on the bandwidth of the filters involved. Thus, Fig. 13.1*j* shows the result of analysing the same signal analysed in Fig. 13.1*i* but with filters accepting a band of frequencies 100 Hz wide at each frequency. This means that the individual harmonics (having a spacing of the voice fundamental of 100 Hz) are smoothed out, but the concentrations of energy at the formant frequencies are still clearly visible. The ear's analysis for frequencies above about 1 kHz is similarly limited (see §15.1). Below 1 kHz, however, the ear's filters are sufficiently narrow that some of the harmonic structure can be resolved.

13.2. BASIC PSYCHOPHYSICAL PARAMETERS OF SOUND

Thresholds

Within the frequency range to which the ear is sensitive, a single sound will always be heard if its intensity exceeds a certain level. Below a somewhat lower level, it cannot be discriminated from the background (or internally generated) noise. Between these levels lies an intensity arbitrarily designated '*threshold*': for example, it is commonly taken as the level at which 50% of the sound presentations are correctly identified. In *Clinical audiometry*, the measurements of

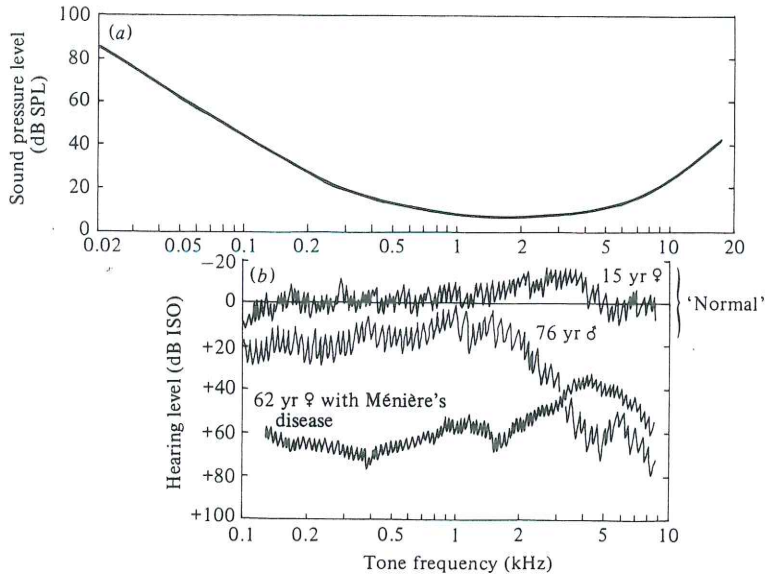


Fig. 13.2. The audiogram, or pure-tone threshold as a function of frequency. (*a*). Pure tone threshold expressed in terms of sound pressure level (dB SPL) at the ear canal under earphone stimulation of one ear. This *minimum audible pressure* (MAP) curve demonstrates the absolute sensitivity of the ear, optimum at the speech frequencies, 0.5–5 kHz. (*b*). Clinical audiograms of 3 subjects by Békésy tracking technique. Each zigzag curve is a tracing of the excursions of tone level resulting from the subject 'tracking' his threshold as the tone frequency is continuously changed, i.e.: he repeatedly increases and decreases the signal level to keep it within the limits of 'just heard' to 'just not heard'. Conventionally, the pure-tone threshold is plotted relative to an internationally accepted standard representing the average hearing of a healthy young adult population (0 dB on scale). Hearing loss appears as a downwards displacement away from zero, as in the case of the patient with Ménières disease. Even in the absence of overt pathology, increasing hearing loss occurs at the high frequencies, with increasing age, as in the case of the 76 year old.

threshold at different frequencies are carried out either by a skilled tester noting the proportion of correct responses at different levels, or automatically by the subject 'tracking' his own threshold, a technique known as Békésy threshold tracking (Fig. 13.2*b*). The threshold so determined, in the quiet, is termed the *absolute threshold.* It can be measured with earphone stimulation of one ear as in Fig. 13.2*a* in direct terms of SPL at the ear canal (i.e. dB with reference

to 20 μPa) or, as is conventional for clinical purposes, relative to an internationally agreed average threshold for young persons free from ear disease (Fig. 13.2*b*). In the latter case, the audiometer compensates for the variation with frequency of the average threshold, and for the characteristics of the earphone, to produce for a person with 'normal hearing' threshold values close to zero (i.e. grouped about the abscissa of the plot). Threshold elevations greater than 10–20 dB are considered abnormal. In audiometry, these elevations in threshold are conventionally plotted *downwards* in dB HL (called 'hearing level' or 'hearing loss') as in Fig. 13.2*b*.

It is clear from Fig. 13.2*a* that the absolute threshold is not constant across frequency. In fact, each species (and indeed each individual) has its own threshold sensitivity curve: the sensitivity is least at the extremes of audible frequency, and maximal at some intermediate frequency. In man, the latter occurs at about 0.5–5 kHz, corresponding to the important frequencies for speech perception. (The most sensitive frequency depends on the dimensions of the ear canal and therefore head; thus, in smaller animals such as cats and guinea-pigs, the most sensitive frequencies are higher: 8–10 kHz (see Fig. 14.5.)

The upper frequency limit of hearing in normal persons is highly variable, approaching 20 kHz in young persons, and diminishing with age as indicated by the audiogram of the 76-year-old man in Fig. 13.2*b* (the latter loss is called presbyacusis; cf. presbyopia, Chapter 4). It has been said that above 20 years, the limit diminishes by about 1 Hz per day!

Loudness

The chief physical correlate of loudness is the level of the sound above threshold. (It is not, however, the only factor, as will be indicated in Chapter 15.)

The subjective magnitude of a sound can be quantified in at least two ways. The first is by *matching* the loudness of the test sound against that of a reference sound, for example a pure tone at 1 kHz. Fig. 13.3 shows the result, averaged across many subjects, of carrying out these matches for tones of different frequencies in intervals of 10 dB SPL. The *equal loudness contours* thus obtained are arbitrarily designated to have the same number of *phons* as the SPL of the matching 1 kHz reference tone. Thus a 60 Hz tone at 60 dB SPL is judged to have the same loudness as a 1 kHz tone at 40 dB SPL and is therefore said to have a loudness level of 40 phons. Loudness

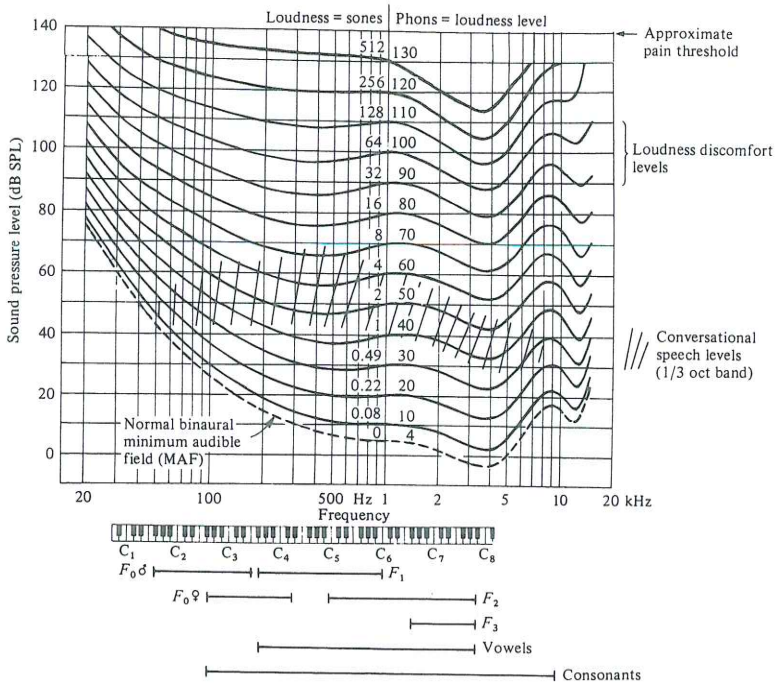


Fig. 13.3. Important subjective functions versus frequency and level. The piano keyboard relates the frequency scale to the musical scale (C_4 is 'middle C'; $A_4 = 440$ Hz). Curve MAF is the internationally agreed average *minimum audible field*, i.e, the threshold curve measured with the subject facing a loudspeaker in the 'free-field' and averaged across many normal young listeners. Because in this case the sound pressure is measured with a microphone *in place of* the listener's head, the MAF differs from the MAP of Fig. 13.2*a* by virtue of the use of two ears, the diffraction of the sound field by the head, and the resonance of the ear canal. The last two effects combine to enhance further the sensitivity to the speech frequencies, particularly near 4 kHz. The curves above the MAF are the internationally agreed contours of *equal loudness*. Each contour is designated the number of phons equal to the sound pressure level (dB SPL) at 1 kHz. The 40 dB phon contour is arbitrarily assigned to have a loudness (= subjective intensity) of 1 sone. Increase in level by 10 dB doubles the loudness. (This simple relationship does not hold below 40 phon.)

The sound pressure levels of the frequencies in conversational speech measured in 1/3 octave bands at about 0.3–3 m from the speaker are indicated approximately by the shaded area. The frequency ranges covered by the voice laryngeal vibrations (fundamental, *Fo*), and the vocal tract resonances (formants: F_1, F_2, F_3) are shown by the lowest horizontal bars. (Equal loudness contours after ISO R 226).

matching ('balancing') techniques are routinely employed in the differential diagnosis of sensorineural hearing loss (Chapter 15, §15.2).

At sufficiently high sound levels, the auditory sensation becomes uncomfortable. This 'discomfort threshold', called the *loudness discomfort level* (LDL), is indicated in Fig. 13.3 at about 100 dB SPL. Above that is the threshold for pain, at 130–140 dB SPL.

The phon, however, whilst used to represent loudness 'level', does not indicate the subjective intensity of a sound. Measurement of this attribute can be carried out by *loudness scaling* techniques in which the subject assigns a number to the perceived magnitude of the sound. These methods give a *power law* relation between the loudness of a sound and its physical intensity, except near threshold. As the exponent of the power law is close to 0.3 for sound levels more than 30 dB above threshold, the loudness function can be simplified to $S = kI^{0.3}$, where S is the loudness estimate, I is the intensity and k is a constant. This means that increasing the sound pressure level by 10 dB leads on average to a doubling of a signal's loudness. It has been internationally agreed to adopt the unit of 1 *sone* as the loudness corresponding to 40 phon. A 50-phon tone sounds on average twice as loud as a 40-phon sound and therefore has a loudness value of 2 sones. The equal loudness contours above 40 phon in Fig. 13.3 have also been labelled in sones.

With headphone listening, the loudness of a signal presented binaurally is almost exactly the sum of the loudness perceived by each ear alone, i.e. it is equivalent to an increase in monaural signal level of about 10 dB.

These scales of loudness and loudness level, determined for the *average subject*, must not be confused with the scale most commonly used in psychoacoustics – the *sensation level* (SL). This simply denotes the physical magnitude (in dB) of a signal above its absolute threshold *for a given individual subject*.

Pitch and timbre

Pitch is the chief psychological correlate of frequency. For pure tones, the ear can locate a signal on a monotonic scale according to its frequency from 'low' (bass) to 'high' (treble) as in Fig. 13.3. The situation is vastly more complicated than this, however, as will be outlined later (§15.1).

'Timbre' is the name given to that quality which distinguishes two steady signals having the same pitch and loudness. It refers to the spectral complexity of the signal, and hence its perceived 'richness',

'mellowness', 'brightness' and so on. Together with the 'attack' (or onset transients of the sound), it allows one to distinguish different instruments of the orchestra playing the same note.

13.3. SUGGESTIONS FOR FURTHER READING

General introductions and reviews

Yost, W. A. & Nielsen, D. W. (1977). *Fundamentals of Hearing*. New York: Holt, Rinehart & Winston.

Durrant, J. D. & Lovrinic, J. H. (1977). *Bases of Hearing Science*. Baltimore: Williams & Wilkins Co. (Physiology and psychophysics, with emphasis on the physics of sound.)

Special references

Sound measurement. Hewlett Packard Acoustics Handbook. Application Note 100. Burns, W. (1973) *Noise and Man*. London: John Murray.

Fourier and signal analysis. Licklider, J. C. R. (1951) Basic correlates of the auditory stimulus, in *Handbook of Experimental Psychology*, ed. S. S. Stephens. New York: John Wiley. Bogert, B. P. (1972) Practising digital spectrum analysis, in *Human Communication: a unified view*, ed. E. E. David & P. B. Denes. New York: McGraw-Hill.

CHAPTER FOURTEEN

Functional anatomy of the auditory system

E. F. EVANS

In this chapter, we consider each of the major elements of the auditory pathway in turn, and summarise its known response properties in terms of its structure. This survey will enable us in the next chapter to account, as far as possible, for the psychoacoustic properties of the auditory system.

Arbitrarily, but conveniently, the auditory pathway is subdivided into the peripheral auditory system, comprising the ear and the primary neurons (auditory or cochlear nerve), and the central auditory system, comprising the nervous pathways and nuclei from the cochlear nuclei onwards (Figs. 14.1, 14.2). The ear itself is conventionally subdivided into the outer ear (pinna and external auditory meatus), middle ear (tympanic membrane, ossicles and associated structures) and inner ear (cochlea), as in Fig. 14.3.

In broad outline, the peripheral auditory system can be considered as a signal 'conditioning' system and spectral analyser (Fig. 14.1). The mechanics of the outer, middle and inner ears transform incident air vibrations into pressure variations in fluids of the inner ear, favouring those frequencies that are important for the organism. In the inner ear, the pressure variations form the input to a bank of filters represented by the mechanical, receptor, and neuronal elements of the organ of Corti.

The central auditory system can in part be considered as a set of parallel systems, parallel in at least two senses. First, signals of differing frequency are represented, in principle, in more or less independent neural 'planes', stacked in the vertical dimension in Fig. 14.1. Secondly, these in turn can be divided into parallel pathways, probably representing different processing subsystems, only two of which, DCN and VCN, are indicated in the plane normal to Fig. 14.1. The former is probably concerned with analysis of the *nature* of a stimulus; the latter its *location* in space.

The diagrammatic layout of these systems seen in Fig. 14.2 is therefore a gross oversimplification of the 'wiring diagram' of the auditory system, expressing ignorance of the detailed anatomy and

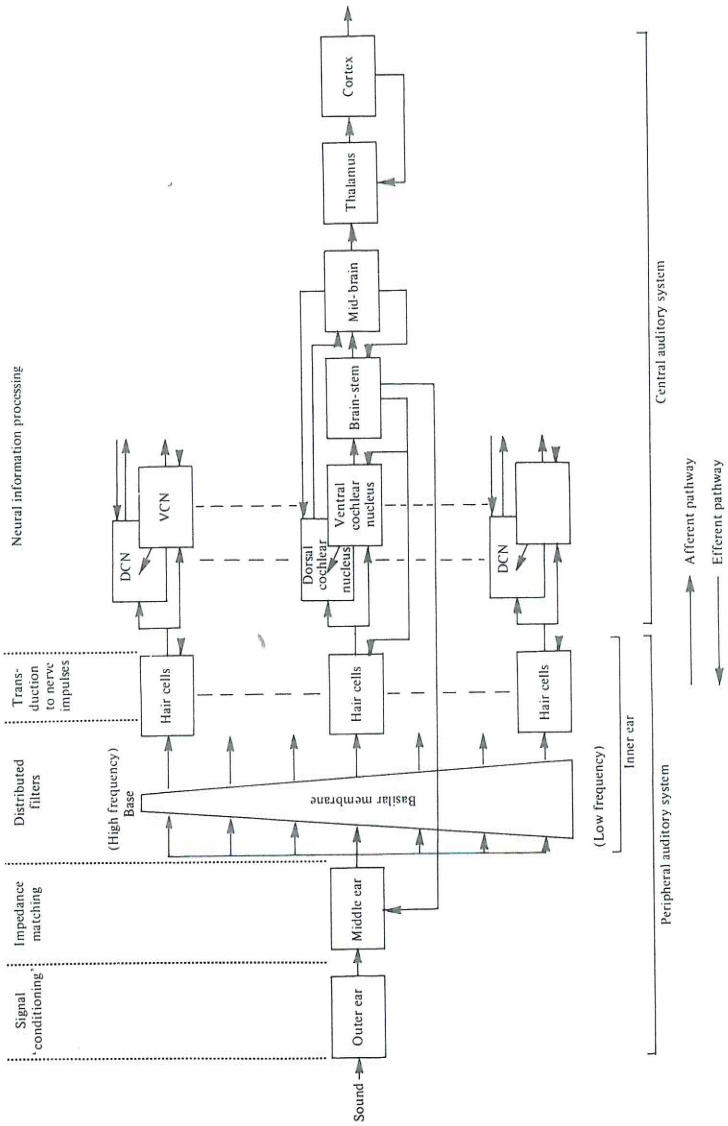


Fig. 14.1. Schematic anatomical and functional map of the auditory system. At the level of the basilar membrane, the system becomes distributed in a spatial dimension (vertical in the figure) that represents frequency. This tonotopic or cochleotopic (see text) organisation is projected to the cochlear nuclei, where further divergence occurs into more than the two subdivisions shown. The cochleotopic dimension is maintained in the neural organisation of the subsequent neural centres. Running in the opposite direction to the afferent pathway (→) is the descending or efferent pathway (←).

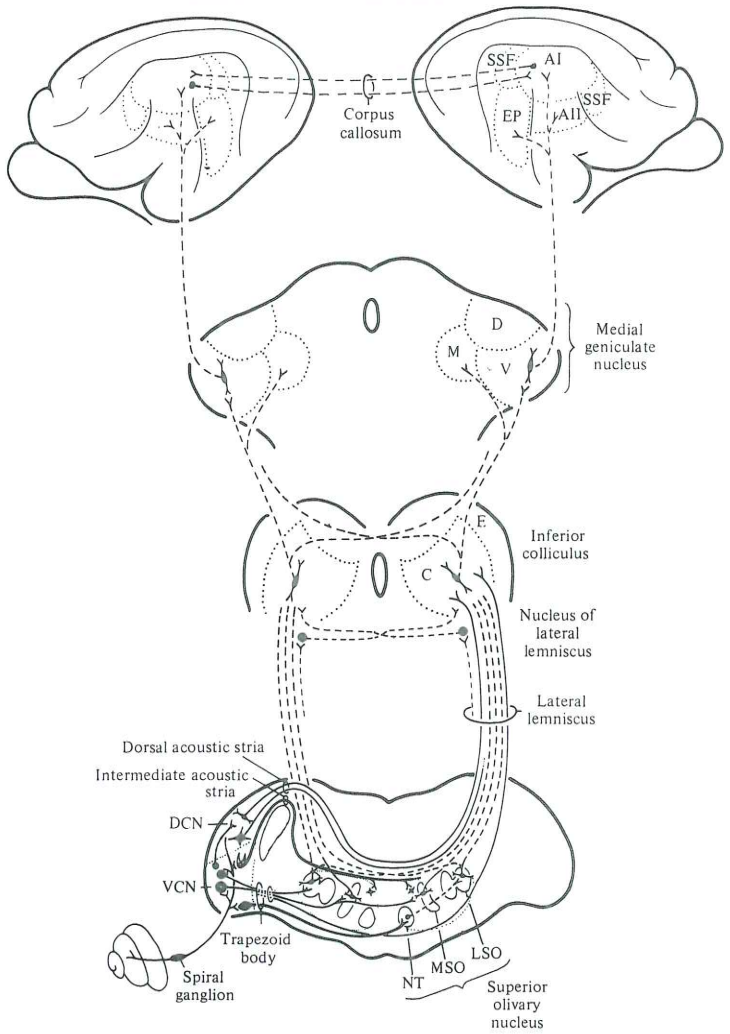


Fig. 14.2. Main ascending anatomical pathways of auditory system. A simplified diagram of the ipsilateral and contralateral projections of one cochlea (bottom left) to the left and right auditory cortex (top) in the cat. DCN: dorsal cochlear nucleus; VCN: ventral cochlear nucleus; NT: nucleus of trapezoid body; MSO/LSO: medial/lateral superior olive; C/E: central and external nuclei of inferior colliculus; M/D/V: medial, dorsal and ventral nuclei of medial geniculate nucleus; AI/AII/SSF/EP: primary, secondary, suprasylvian fringe and posterior ectosylvian areas of auditory cortex. (Modified, after Konigsmark, *Archives of Otolaryngology*, *98*, 403.)

function but serving usefully as a basis for description. Furthermore, running as a 'counter-current' to the afferent auditory system, is the *efferent* system (Fig. 14.1). Regrettably, we have little information on the functional interrelationships of these two systems.

Most of our detailed knowledge of the anatomy and physiology of the auditory system comes from studies on animals, particularly the cat. We have far more data available on the peripheral than on the central auditory system, because of the former's relative accessibility and simplicity of structure and function. This is not too unfortunate, for lesions of the peripheral auditory system are the most common cause of hearing loss.

14.1 OUTER EAR

The *auricle* or *pinna*, and the *external auditory meatus* or *canal* together constitute the outer ear (Fig. 14.3). The canal is relatively straight in man, having a diameter of about 0.7 cm and a length of 2–3 cm.

The convolutions of the pinna are of considerable importance for the perception of sound direction both horizontally and vertically and as coming from outside the head.

The shape of the head, pinna and ear canal together modify a plane sound field in such a way that the pressure at the tympanic membrane shows a broad but substantial gain of 10–20 dB at the important frequencies for speech signals: 2–5 kHz. This frequency-dependent gain is much reduced when headphones are used, and is almost lost in the case of the 'insert' ear-pieces of hearing aids.

14.2. MIDDLE EAR

The middle-ear structures are located in an air-filled cavity of the temporal bone, vented to atmosphere via the *eustachian tube* (Fig. 14.3). The cavity is bounded laterally by the cone-shaped *tympanic*

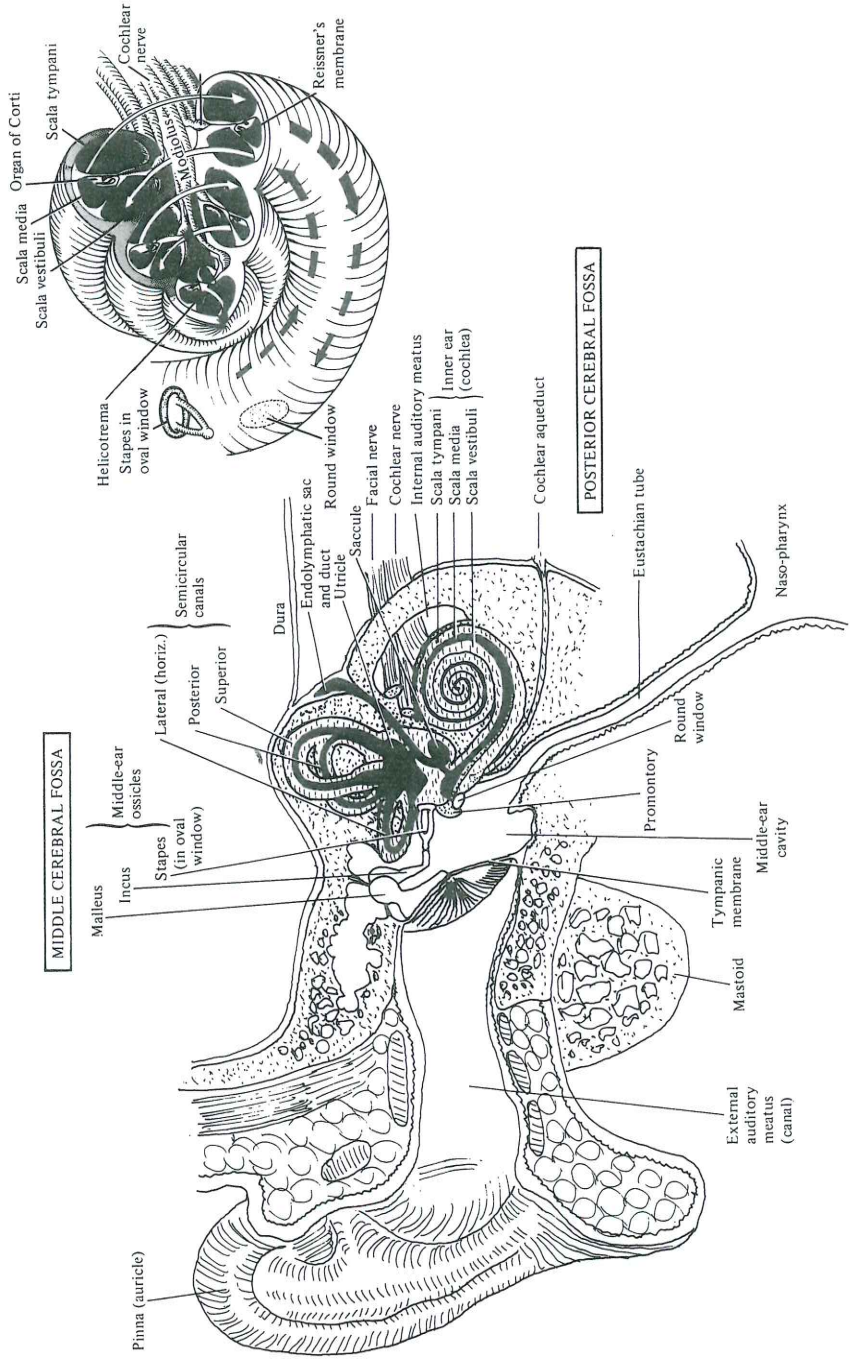


Fig. 14.3. Semi-diagrammatic cross-section of right outer, middle and inner ear of man as viewed from in front. Bone showed as stippled areas; perilymph in osseous labyrinth, dashed areas; endolymph in membranous labyrinth, black areas. Inset shows more precisely the orientation of the cochlea, with its apex laterally and slightly anteriorly. The arrows indicate the continuity of the scalae vestibuli from the stapes-filled oval window via the helicotrema to the scala tympani and the round window. (Inset from Curtis, Jacobson & Markus (1972) *An Introduction to the Neurosciences.* W. B. Saunders Co.)

membrane or *ear drum*, to the upper vertical radius of which is attached the elongated *manubrium* of the first of the *ossicles*, the *malleus* ('hammer'; see Fig. 14.4). The malleus pivots about a horizontal axis through the firm junction between its almost spherical head and the *incus* ('anvil'). Rotations of the malleus—incus combination are transmitted, via the long process of the incus and the *incudostapedial joint*, to the head of the *stapes* ('stirrup'). The two *crura* of the stapes transmit movements to the stapes *footplate*, which fits into the membrane-covered *oval window* of the cochlea and acts like a piston therein. The ossicles of the middle ear are suspended by ligaments and are acted upon by two muscles (Fig. 14.4) the *m. tensor tympani*, attached to the manubrium of the malleus and the *m. stapedius* inserted into the neck of the stapes.

The function of the middle-ear system is to transform efficiently variations in air pressure in the ear canal into pressure variations in the fluids of the inner ear. Since the impedance* of air is very low compared with that of the cochlea (about 1:135), if the energy were presented directly via an air–fluid interface, 97% would be reflected, and only 3% transmitted. Because of the impedance-matching function of the middle ear, an estimated 60% of incident energy is in fact transmitted into the cochlea. This *impedance transformer* operates predominantly by virtue of the substantial difference between the effective area of the tympanic membrane and that of the footplate of the stapes (areal ratio of 17:1). In addition, because the length of the long process of the incus is somewhat shorter than that of the malleus, there is a small mechanical advantage of 1.3:1 in terms of lever ratio. The overall pressure ratio is therefore, in principle, about 1:22 in man. (This represents an impedance ratio of about 1:29.) In practice, this pressure gain is approached only at frequencies intermediate between about 1 kHz and a few kHz in man.† The

* Acoustic impedance at any frequency is the vectorial sum of *reactance* (determined by the opposing effects of mass and stiffness (elasticity) of the system) and *resistance* (produced by friction). It is measured in acoustic ohms. The inverse of acoustic impedance is *admittance*, measured in mhos. The inverse of the stiffness component is *compliance*, i.e.: compressibility. Thus air, being very substantially more compressible than a fluid, has a much lower impedance.

† About 20 kHz in cat and guinea-pig. In this range, the impedance is minimal, and chiefly resistive. For frequencies below this range, the stiffness of the middle-ear and inner-ear membranes (including the tympanic membrane), the middle-ear ligaments and the air in the middle-ear cavities, dominate the impedance, together with a shunting effect of the helicotrema at very low frequencies. For frequencies above the range, the mass component dominates the impedance; this together with reduction of the effective area of the tympanic membrane increasingly limits the transfer.

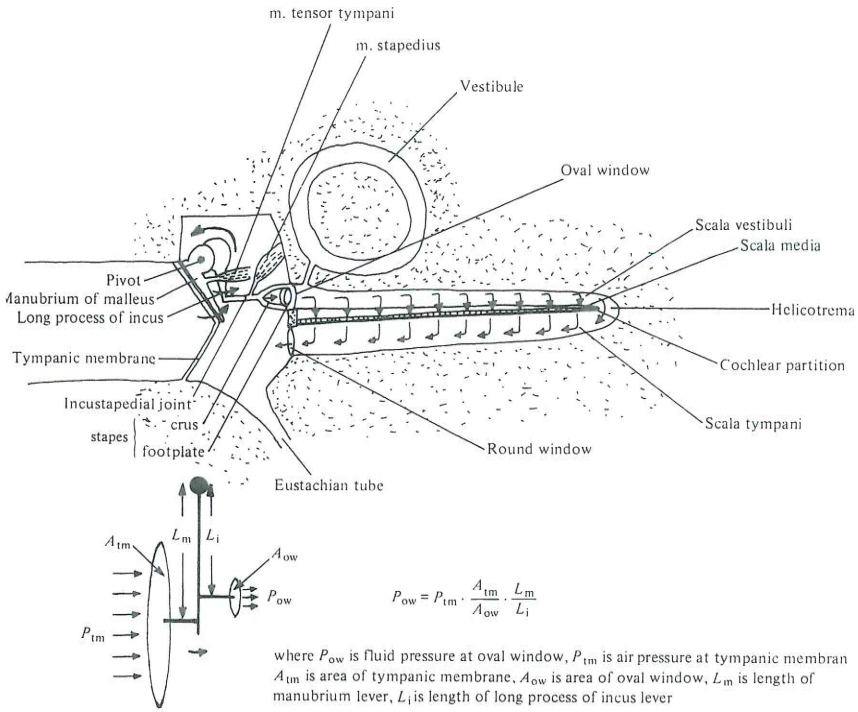


Fig. 14.4. Schematic functional diagram of middle ear and uncoiled inner ear, and middle ear transformer. Vibrations of tympanic membrane are transmitted as rotations of malleus and incus about common axis (normal to page, marked with dot). This produces piston-like movements of the stapes footplate in the oval window, with transmission of the pressure changes across the cochlear partition virtually instantaneously throughout the cochlear length. Arrows indicate direction of movements in response to a compression wave. The lower diagram illustrates transformer action of the middle ear by virtue of the large difference in area of tympanic membrane and oval window, and the (smaller) lever ratio (L_m/L_i). The middle ear muscles are contained in bony canals. The m. stapedius acts sideways on the incustapedial joint to stiffen the ossicular chain.

middle ear therefore, like the outer ear, emphasises intermediate frequencies. The combination of the two almost entirely accounts for the form of the threshold audiogram (Fig. 14.5).

The input impedance of the middle ear can be measured for clinical purposes by determining the attenuation of a low-frequency probe tone, introduced into the sealed external auditory canal from a source

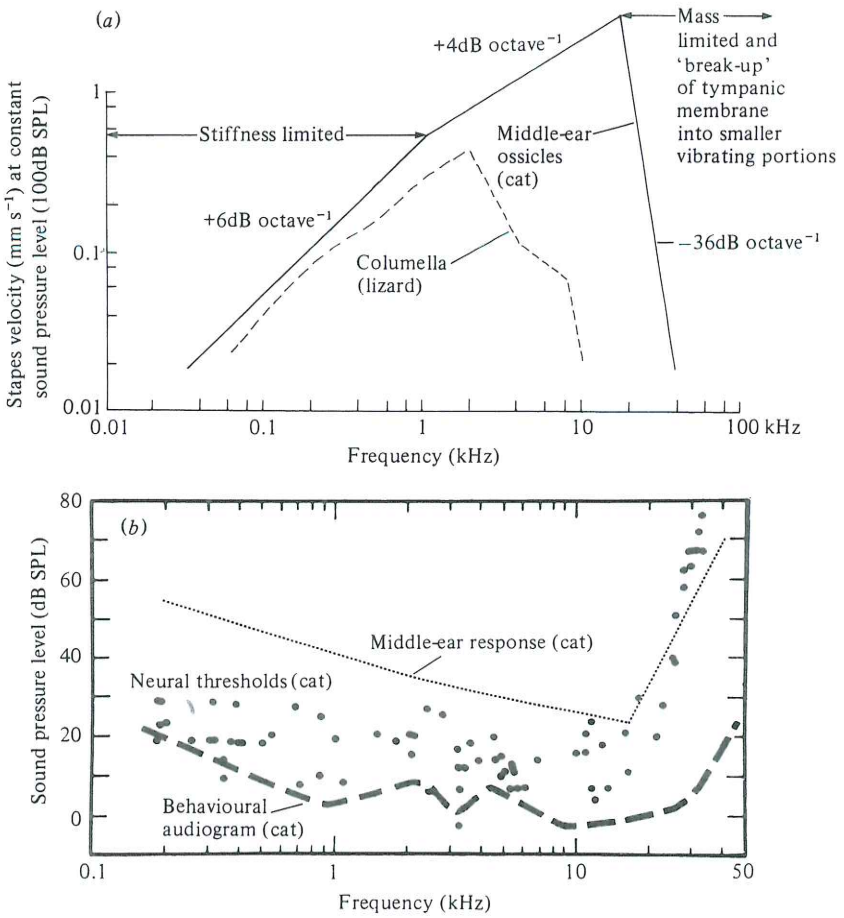


Fig. 14.5. Contribution of middle ear to ear's sensitivity. (*a*). Frequency response of cat's middle ear (continuous line) measured in terms of velocity of stapes motion at constant SPL at the tympanic membrane. (Velocity is used rather than displacement because it represents pressure in the inner ear.) Analogous response in lizard (dashed lines) has much less extended frequency response, probably because the latter possesses only a *columella*, a single peg-like ossicle linking tympanic membrane and inner ear, and not the mammalian ossicular system. (Lizard response from Johnstone & Sellick (1972) *Quarterly Review of Biophysics*, *5*, 1; cat response from Wilson & Evans, unpublished data.) (*b*) Comparison between behavioural (audiogram) and neural thresholds of cat and its middle-ear response. (The vertical position of the middle-ear response is arbitrary.) Each point represents the threshold of an individual cochlear fibre at its characteristic frequency. Much of the ear's sensitivity is therefore accounted for by the middle-ear response. The resonance of the ear canal is responsible for most of the remaining differences, particularly between 1 and 10 kHz. (Audiogram after Neff & Hind (1955) *Journal of the Acoustical Society of America*, *27*, 480, corrected for outer-ear and bulla response; neural thresholds from Kiang (1968) *Annals of Otology*, *77*, 656. After Evans (1975) Cochlear nerve and cochlear nucleus in *Handbook of Sensory Physiology*, vol. V/2, Chapter 1, Springer-Verlag.)

of known impedance. It is commonly expressed as the reciprocal of the stiffness component, i.e. as compliance, usually in terms of the volume of air having the same compliance. By subtracting out the compliance of the external canal, the compliance of the middle-ear–inner-ear system itself can be estimated. This provides information on the mobility of the ossicular chain, and is a valuable measure clinically, particularly when combined with information on the changes in ear-canal compliance with canal pressure, measurements known as *tympanometry*. Deviations in compliance much above the normal range (of about 0.3–2.5 ml) indicate abnormal mobility of the drum, as would be produced by traumatic interruption of the ossicular chain. Conversely, abnormally low compliance indicates a reduction in the mobility of the middle ear, either from fixation of the stapes footplate as in otosclerosis, or from abnormally high or low pressure in the middle ear (as with a blocked eustachian tube), or from fluid in contact with the tympanic membrane.

These deficiencies in middle-ear transmission produce what is known as *conductive hearing loss*. These losses rarely exceed 60–70 dB because of direct conduction of sound incident upon the head through the skull to the cochlea. This *bone conduction* affords another simple test for lesions of the middle ear. In addition to the audiogram for airborne sound (Fig. 13.2), a bone-conduction audiogram is determined by using a calibrated vibrator pressed to the mastoid bone. The calibration is such that, in the absence of middle-ear disease, the 'bone' and 'air' audiograms coincide. In the presence of middle-ear disease, a 'bone–air gap' appears between the two audiograms, the magnitude of the gap indicating the magnitude of the attenuation in middle-ear conduction.

Transmission of energy through the middle ear can be modified by the middle-ear muscles (Fig. 14.4). The tensor tympani (innervated by the Vth cranial nerve) pulls on the manubrium to increase (as its name indicates) the tension of the tympanic membrane; the stapedius

(innervated by the VIIth nerve) pulls on the neck of the stapes at right angles to its 'piston' axis, thereby tending to immobilise the footplate. Both actions increase the stiffness of the middle-ear system, and so attenuate the transmission of energy at the lower frequencies particularly below about 2 kHz, by a factor approaching 30 dB at 100 Hz. The muscles of the middle ear (mainly or only the stapedius) contract *reflexly* in response to high level (> 80 dB SL) sounds received *by either ear* with a latency of effect on middle-ear transmission of 15–150 ms and about 100–500 ms for maximum action, depending upon stimulus level. Note that the reflex can be evoked in one ear by acoustic stimulation of the other. This property is useful for diagnostic purposes.

The reflex can protect the inner ear against continuous intense sounds. Attenuations of 0.6–1 dB for every dB increase in acoustic input have been measured, directly in animals and indirectly in man, for SLs above 100 dB. More significantly, the middle-ear muscles are contracted just prior to and during vocalisation in order to reduce self-stimulation. In man, the preferential attenuation of lower frequencies means that the acoustic reflex would be effective in reducing the masking of high-frequency signals by low-frequency sounds. This is of potential importance for reducing the masking caused by the high-energy low-frequency components of one's own voice on the less intense higher-frequency components (e.g. in consonants) of simultaneously received speech from another speaker. The threshold and time course of the middle-ear muscle reflex, elicited by tone pips, can be determined by dynamic impedance measurements as above, and is clinically useful in evaluating ossicular function and the integrity of the cochlear nerve and brainstem nuclei.

14.3. INNER EAR

Morphology and biochemistry

The inner ear is an extremely delicate organ buried, in man, in the hardest bone in the body, the *petrous temporal bone* (Fig. 14.3). It consists of a series of passages, the *osseous labyrinth*, containing a complex system of sacs and tubes, the *membranous labyrinth* (which appears black in Fig. 14.3) surrounded by a fluid known as *perilymph*. The membranous labyrinth houses the receptor cells of the vestibular and cochlear systems, and contains a fluid of different composition from the perilymph, the *endolymph*.

The cochlear component of the labyrinth is coiled like a snail shell (hence the name *cochlea*), around a central axis, the *modiolus*, occupied by the nerve trunks. It has nearly three turns in man. Its base lies at the oval and round windows, where it connects with the vestibular labyrinth, and its apex lies more anteriorly and laterally than appears in Fig. 14.3 (see inset). The membranous labyrinth partitions the coiled labyrinth into three channels, shown diagrammatically in Figs. 14.3 and 14.4, and in cross-section in Fig. 14.6. The more anterior channel, the *scala vestibuli*, communicates with the vestibular labyrinth (Fig. 14.3) and hence the oval window, and therefore receives the acoustic input to the cochlea (Fig. 14.4). The more posterior channel, *scala tympani*, terminates in the membrane-covered *round window*, which acts to release intracochlear pressure.

The middle channel, the *scala media*, or *cochlear duct*, contains the acoustic sensory epithelium, the *organ of Corti* supported by the *basilar membrane* which, with the spiral lamina, forms the *cochlear partition*. The length of the basilar membrane is about 34 mm in man (compared with about 20 mm in cat and guinea-pig). It tapers in width from about 500 μm at the apical (helicotrema) end to about 100 μm at the basal end where, conversely, the scalae are largest. Here, the osseous spiral lamina spans most of the scalar width.

The organ of Corti in mammals is about 100 μm in thickness and consists of two types of sensory *hair cells* rigidly attached to the basilar membrane by supporting cells (*Deiter's cells*) and the *pillars* (Fig. 14.6). The hair-bearing ends of the hair cells are held firmly together in the rigid *reticular lamina*, continuous with the heads of the pillars. There is a single row of *inner hair cells* and 3–5 rows of *outer hair cells* (Figs. 14.6, 14.7) each row containing about 100–130 hair cells mm^{-1}. The hair cells bear stiff sensory hairs or *stereocilia* (3–6 μm in length), composed of filaments of actin. The stereocilia are arranged in characteristically shaped rows: in a shallow U formation, for the inner, and in a V or W formation for the outer hair cells (Fig. 14.7). The cilia of the outer hair cells are embedded in the *tectorial membrane*, a gelatinous structure overlying the hair cells and anchored to the outermost supporting cells by a marginal net (Fig. 14.6). Those of the inner hair cells appear not to be embedded at all in some species (guinea-pig, primates) or not firmly (cat).

As indicated in Fig. 14.3, the perilymph-filled spaces of the labyrinth communicate with the posterior cranial fossa via the *cochlear aqueduct* leading to the scala tympani near the round

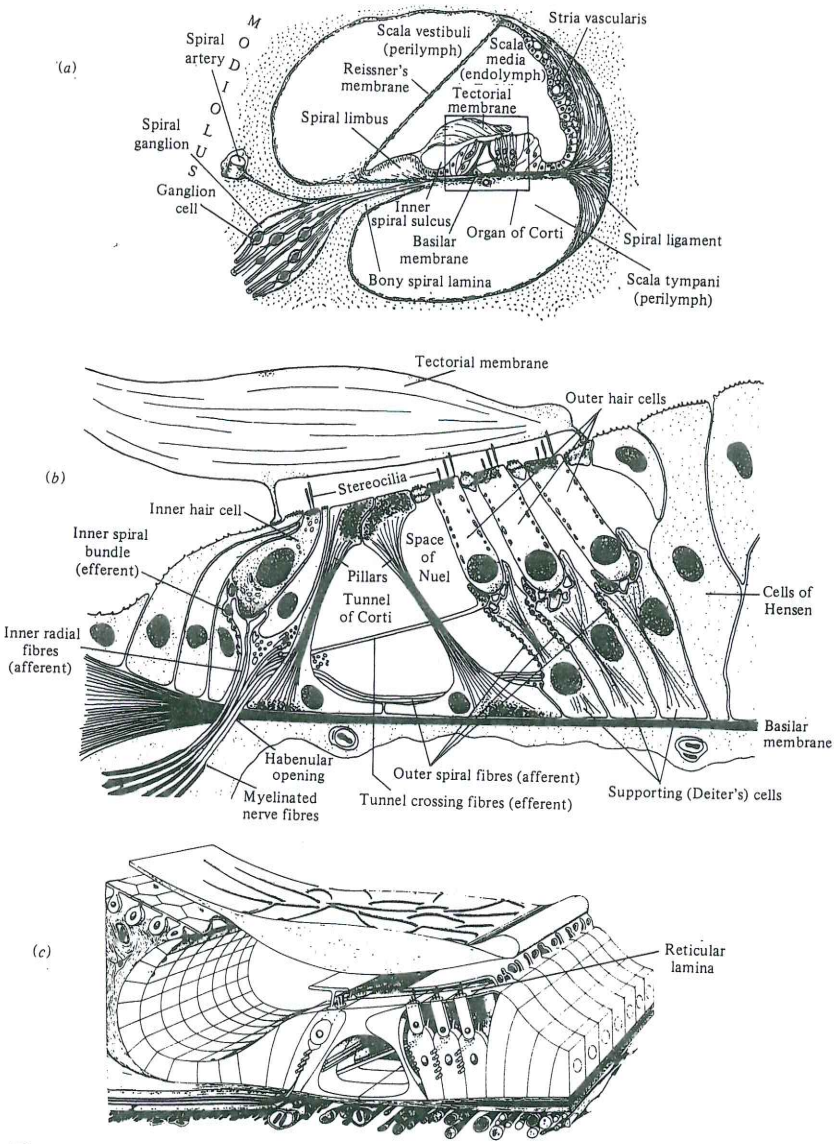


Fig. 14.6. Detail of cochlear partition and organ of Corti. (*a*) Cross-section through cochlear duct, modiolus of cochlear spiral being to the left of the diagrams. (For orientation see inset of Fig. 14.3.) (After Rasmussen (1933) *Outlines of neuro-anatomy*. William Brown Co.) (*b*) Enlarged view of organ of Corti. (After Durrant & Lovrinic (1977). *Bases of Hearing Science*. Williams & Wilkins Co.) (*c*) Three-dimensional view of hair cells and tectorial membrane. (After Lim (1972) *Archives of Otolaryngology*, *96*, 199.)

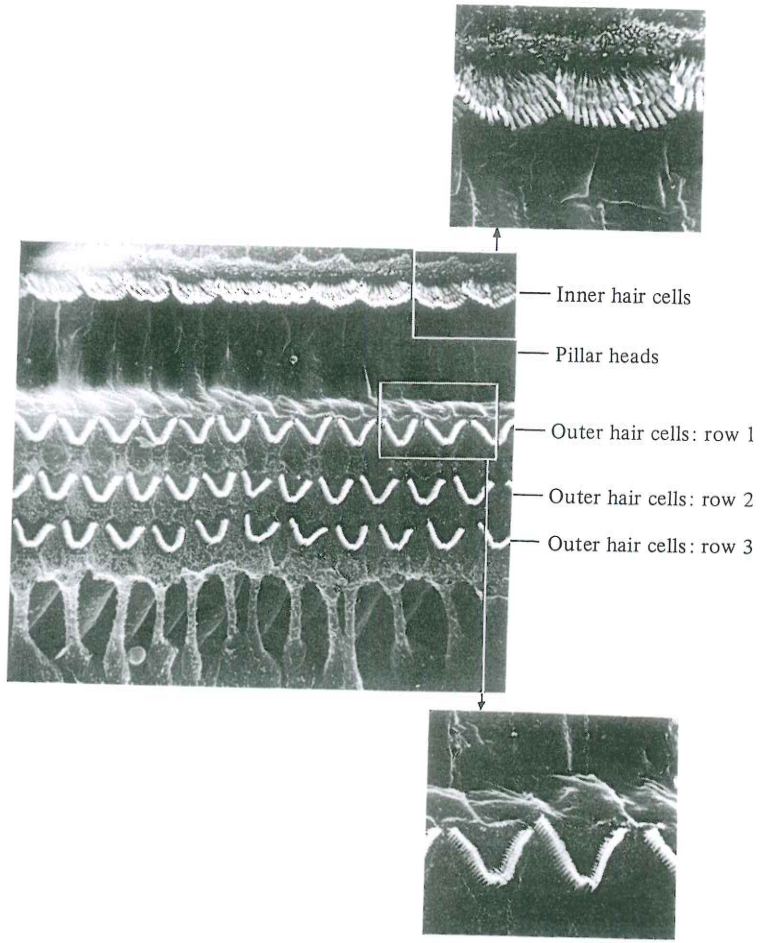


Fig. 14.7. Scanning electron microscope views of the surface of the cochlear hair cells after removal of tectorial membrane. Insets show enlarged views of the cilia (hairs) of the inner and outer hair cells. (From Bredberg *et al.* (1972) *Acta Otolaryngology*, *Supplement*, 301.)

window. The composition of the perilymph is therefore virtually identical to cerebrospinal fluid, i.e. high in sodium ions (about 150 mM l^{-1}) and low in potassium ions (3–4 mM l^{-1}). In contrast, endolymph contains a high concentration of potassium ions (about 150 mM l^{-1}) and a very low concentration of sodium ions (1–2 mM l^{-1}). It is likely that the cells of the *stria vascularis* (Fig. 14.6) contain an electrogenic ion pump, which actively transports potassium

ions into, and sodium ions out of the endolymph, and is responsible for the standing intracochlear potential (see *Cochlear potentials*).

Surrounding the outer hair cells and the outer pillars are a series of spaces which interconnect with that between the pillars, the *tunnel of Corti* (Fig. 14.6). They are filled with *cortilymph*. This probably has a similar composition to the perilymph. Thus, with the exception of their hair-bearing ends, the hair cells have a similar extracellular milieu to that of cells elsewhere in the body. On the other hand, the high concentration of potassium ions in the region of the apical (hair bearing) ends of the hair cells may be essential to mechanoreception. It is also found in the lateral-line systems of fishes and amphibia.

The blood supply to the labyrinth (from the basilar artery directly, or via the anterior inferior cerebellar artery) enters the cochlea by way of the internal auditory meatus. It is therefore easily occluded, for example by tumours of the VIIIth nerve. There are no blood vessels in the organ of Corti itself, but a spiral artery shown in Fig. 14.6 runs along the basilar membrane in some species including man. The organ of Corti is extremely sensitive to reduction in its oxygen supply, which comes mainly from capillary networks in the vicinity of the spiral lamina, under the tunnel of Corti, and in the spiral limbus.

Cochlear mechanics

The piston-like movements of the stapes footplate in the oval window create fluctuations in pressure in the perilymph of the scala vestibuli (Fig. 14.4). These are transmitted with virtually no delay throughout the scala vestibuli and, through Reissners membrane, to the scala media. Because the membrane of the round window lacks stiffness, the pressure in the scala tympani is virtually constant at that of the middle ear. Hence, a temporally varying pressure difference is established across the cochlear partition. Because the stiffness of the basilar membrane changes along its length as the width tapers from the apical to the basal end, these pressure differences set up *travelling waves* in the membrane itself (Fig. 14.8).* These waves

* It is important to point out that a travelling wave motion of the basilar membrane does not necessarily imply transfer of energy along the membrane itself (in the manner of a wave travelling along a 'flicked' rope). The degree of coupling between segments of the basilar membrane along its length is slight, and it is the pressure difference across the length of the cochlear partition that is the driving force for the travelling wave. This means that the travelling wave will propagate from base to apex irrespective of where the pressure changes originate (as in bone conduction via the bones of the skull), and accounts for the fact that travelling waves produced by low frequencies are not impeded by damage to or calcification of the basal part of the cochlear partition.

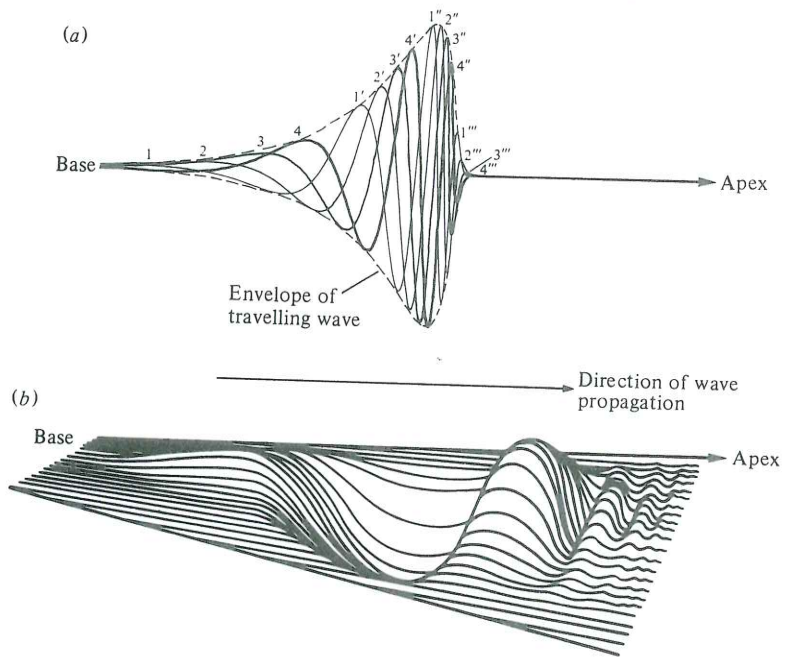


Fig. 14.8. Diagrams of travelling wave motion of the basilar membrane. In both cases, the amplitudes of the motion have been exaggerated for clarity. Even above the threshold of discomfort (120 dB SPL) the amplitude of vibration would be less than 1% of membrane width. (*a*) Two-dimensional representation of travelling wave at instants of time corresponding to 1/3 period (1, 2, 3, 4, 1′, 2′, 3′, 4′, etc.) (Courtesy of G. J. Sutton.) (*b*) Three-dimensional representation of travelling wave in small segment of basilar membrane, at one instant in time. (From Tonndorf (1960) *Journal of Acoustical Society of America*, *32*, 238.) Note wave travelling from base (to left) toward apex (out of view to right), with its *envelope* having a maximum at a given location, and a sharp cut-off towards the apex.

propagate from the basal to apical end, and as they do so, they grow in amplitude gradually to a peak, then rapidly collapse. The position of the peak amplitude depends upon the stimulus frequency, so that a map of peak frequencies can be established for the basilar membrane, the highest peak frequencies being represented at the basal end and the lowest at the helicotrema. The velocity and the wavelength of the travelling wave decrease with distance from the stapes, both decreasing very rapidly beyond the point of peak motion, as does the amplitude of vibration.

At any one point on the basilar membrane, the amplitude of

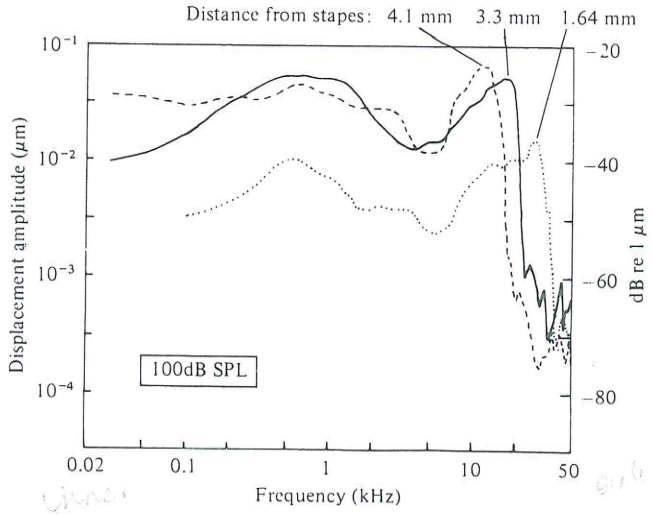


Fig. 14.9. Frequency response of individual basilar membrane locations. Amplitude of vibration *vs.* frequency of three places on guinea-pig basilar membrane, as recorded by capacitive probe *for constant sound pressure at tympanic membrane (100 dB SPL)*. Location of places: 1.64, 3.3 and 4.1 mm from stapes. Each curve represents the amplitude, at a given location, of the envelope of the travelling waves (as in Fig. 14.8) corresponding to the frequencies examined. The smooth curve of Fig. 14.8 is therefore transformed (from amplitude as a function of position at a given frequency) to response as a function of frequency at a given location, and is 'distorted' by the response of the middle ear (pressure at the tympanic membrane being constant across frequency). This produces a low-pass filter characteristic: little change of the vibration amplitude at low frequencies; rapid cut-off at high frequency side of frequency of maximum vibration. (From Wilson & Johnstone (1972) in *Hearing Theory*, p. 172, IPO Eindhoven.)

motion shows broad tuning with frequency (Fig. 14.9). Unfortunately, not all studies are in agreement on the degree of tuning, and on the linearity of vibration.* The majority, however, indicate that the motion of a point on the basilar membrane is more or less *low-pass* in character (Fig. 14.9). This follows from the shape of the envelope of the travelling wave (dashed line in Fig. 14.8): all locations on the basilar membrane can be more easily activated by frequencies lower than their peak frequency than by higher frequencies.

The peak movements of the basilar membrane are some 30 times greater in amplitude than those of the stapes footplate. Even so, they

* Very recent studies suggest that the tuning shown in Fig. 14.9 represents that of the *passive* basilar membrane.

are extremely small, of the order of 0.01–0.1 μm (10–100 nm) at 100 dB SPL. Thus, at threshold (ca. 0 dB SPL), they must be less than atomic dimensions, on the assumption of linearity of vibration!

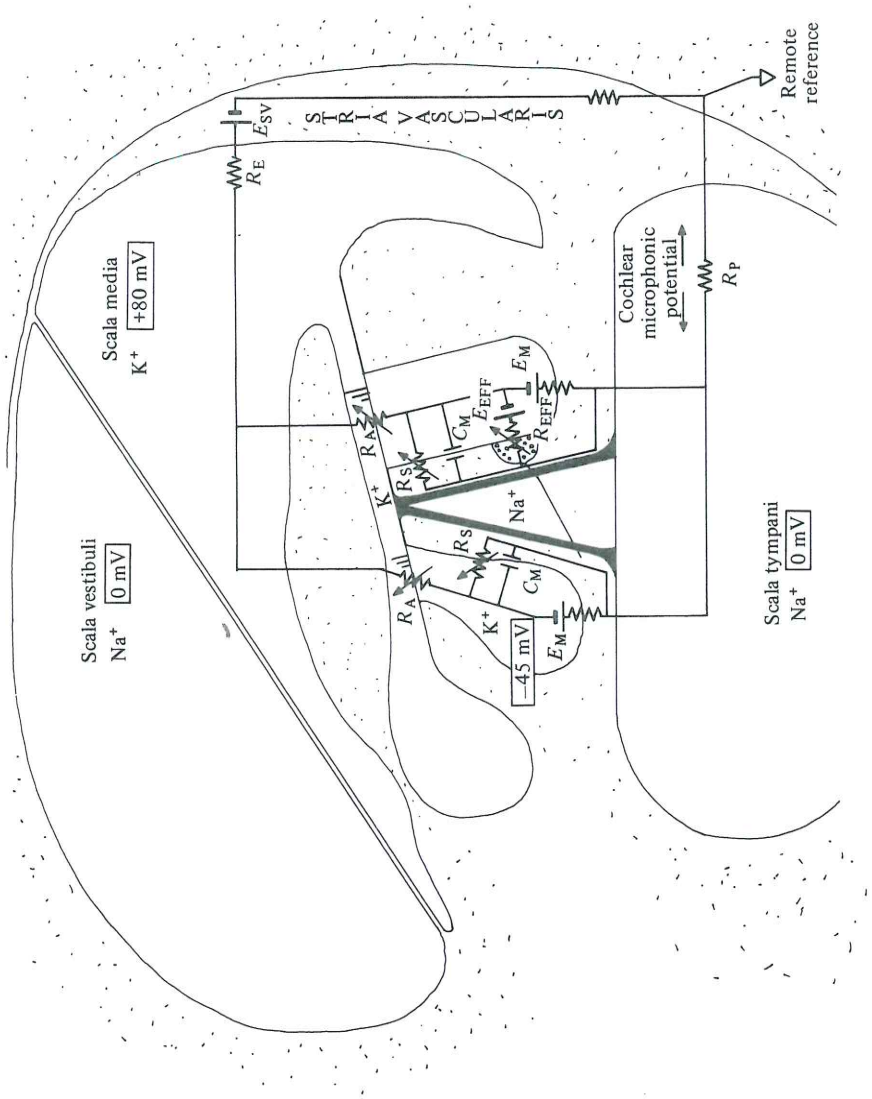
It is generally held that motion of the cochlear partition away from and towards the tectorial membrane induces a shear motion of the hairs of the hair cells in a radial direction (i.e. across the cochlear partition). It must be emphasised, however, that because the movements of the cochlear partition are so small compared with its dimensions, and because the physical properties of the tectorial membrane and the nature of contact between it and the cilia of the inner hair cells are not yet clear, it is as yet impossible to measure the exact motion of the cilia. It is assumed that movements of the stiff cilia are the necessary antecedents of the electrical changes in the hair cells to be described below. However, it is at present impossible to decide how this occurs (see Fig. 14.6): whether direct brushing of the cilia against the tectorial membrane is involved, as is likely for the outer hair cells, or whether, in the case of the inner hair cells, the cilia are displaced by the motion of fluid in the subtectorial space, or whether other, complex forms of 'micromechanics' of tectorial membrane, fluid and hair cells are involved. This is an important question, for it appears that the tuning of the responses of inner hair cells and cochlear nerve fibres is sharper than that of the basilar membrane (see p. 277), and that the presence of a tectorial membrane and outer hair cells is necessary for the sharp tuning.*

Cochlear potentials

Standard potentials. Fig. 14.10 summarises the results of exploration of the cochlea by microelectrodes. The d.c. potential in the endolymph of the scala media (endolymphatic potential) is +80 mV relative to that of the perilymph or a remote electrode. This potential is probably maintained (against a potassium ion diffusion potential in the opposite direction) by the electrogenic pump (E_{SV}) in the stria vascularis. The pump may be that responsible for the transport of potassium ions into and sodium ions out of the endolymph. Abolishing the activity of the pump, by anoxia, leads to a rapid reduction and reversal of the endolymphatic potential.

This 'battery' (E_{SV}) appears to act in conjunction with the transmembrane battery (E_M) of the hair cells, to produce a standing current flow through the hair cells. The latter battery (E_M) is the classical ionic diffusion potential (about −60 mV) responsible for the negative intracellular potential of cells in general, based on the ratio

* Recent evidence emphasises the active role of the organ of Corti, including the basilar membrane, in accounting for sharp cochlear tuning.



of potassium ion (and/or chloride ion) concentrations between the hair cell interior and the cortilymph.

Receptor potentials, transduction, and generation of action potentials. As indicated above, the form of the mechanical stimulus to the hair cells of the mammalian cochlea is not known precisely. However, by analogy with hair cells in vestibular and lateral-line systems, it must involve displacement of the cilia. Intracellular recordings from such hair cells (in the mammalian cochlea so far mainly from the inner hair cells) show that a sufficient stimulus produces a depolarising *receptor potential* in the inner hair cell (Fig. 14.11), accompanied by a decrease in the resistance between the interior and exterior of the cell. The receptor potential is generally a distorted version of the mechanical stimulus waveform (Fig. 14.11*a*). At frequencies higher than about 300 Hz in the mammalian inner hair cell, the low-pass filtering characteristic of the cell membrane (represented by C_M and the resistances of Fig. 14.10) progressively attenuates the a.c. component of the receptor potential, leaving predominantly a depolarising d.c. component for the frequencies above a few kHz (Fig. 14.11v). As will be shown later there is evidence for a cyclical component of excitation of cochlear nerve fibres known as *phase locking*. This is probably attributable to the a.c. component of the receptor potential; it attenuates with frequency above 1 kHz, but is still evident up to 5 kHz.

Fig. 14.10. Simplified diagram of electrical potential generators and current paths in cochlea. E_{SV} electrogenic pump in stria vascularis, responsible for the endolymphatic potential of +80 mV. Displacement of the hair cells in the effective direction probably reduces the series resistance through the apical, hair-bearing end of the hair cells (R_A) or the hair cell transmembrane permeability (shunt resistance, R_S), producing the *depolarising* receptor potential recorded intracellularly (Fig. 14.11) superimposed as the standing resting potential of about −45 mV. This in turn results from the shunting currents through R_A and R_S reducing the effects of the transmembrane diffusion potential battery (E_M, probably about −60 mV). In the case of the outer hair cells, activity of the efferent terminals decreases R_{EFF} hence *hyperpolarises* or stabilises the intracellular hair cell potential.

The receptor currents flow through the low extracellular resistance paths R_E and R_p (endolymph and perilymph) thus generating small extracellular potentials, such as the *cochlear microphonic* potential, recorded in the extracellular spaces and outside the cochlea (Fig. 14.12).

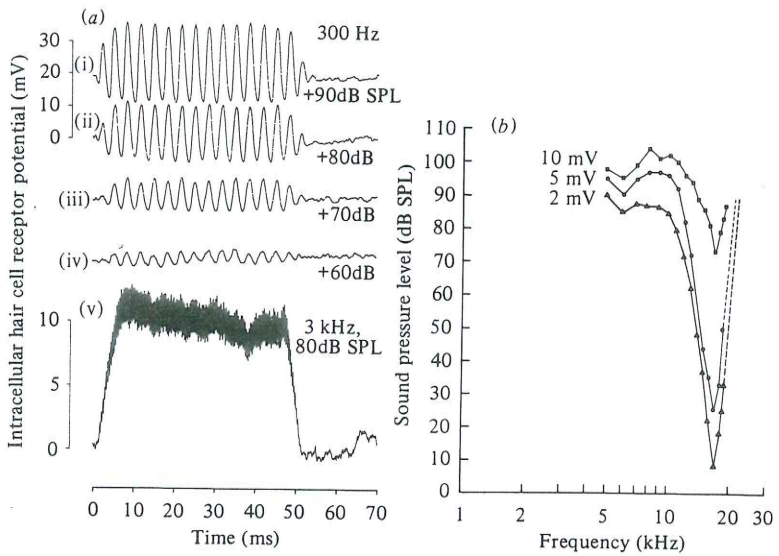


Fig. 14.11. Intracellular potentials from mammalian inner hair cells. (*a*) Intracellular potential waveforms recorded by a microelectrode in an inner hair cell in the basal (high frequency turn) of a guinea-pig: (i)–(iv) to a tone burst of low frequency (300 Hz), and at different sound-pressure levels. The waveform of the stimulus, distorted at high sound levels, can be seen in the response. (v) response to 3 kHz tone. Note small a.c. response superimposed upon a substantial (*c.* 12 mV) depolarising potential. (*b*) Variation in stimulus level required to keep d.c. receptor potential constant across frequency for three voltages (isovoltage frequency tuning curves). ((*a*) courtesy of Merzenich, Russell & Sellick, and (*b*) after Russell & Sellick, (1978) *Journal of Physiology*, *284*, 261.)

The amplitude of the receptor potential in mammalian inner hair cells depends on the intensity and frequency of the sound stimulus. Each hair cell has a most sensitive frequency (17 kHz in Fig. 14.11*b*); at progressively higher intensities, the band of frequencies evoking a response grows progressively wider until at the maximum sound level, frequencies from 1 to 19 kHz evoke a response, and the maximum d.c. receptor potential approaches 20 mV. By choosing a constant response amplitude, an isoresponse intensity–frequency function can be plotted, as shown in Fig. 14.11*b*. This shows how sharply tuned are the receptor potentials of inner hair cells.

In the reptilian cochlea there is evidence that the receptor potential is a component in a resonance mechanism within the hair cells

themselves. This mechanism is possibly ionic or electromechanical in nature and is responsible for the sharp tuning. Whether this is the case also in the mammalian inner hair cells is not known.

By contrast, outer hair cells do not appear to have d.c. receptor potentials; so far only a small (few mV) a.c. voltage has been recorded. This appears to be as broadly tuned as the amplitude of vibration of the basilar membrane, and appears likely to be the origin of the extracellularly recorded cochlear microphonic (see below).

It is not certain at present how the hair cell receptor potential is generated. In the most widely accepted explanation (the Davis model), displacement of the hairs alters the resistance across the *apical* end of the hair cell (R_A in Fig. 14.10) and thus alters the standing current, driven through the hair cell by the two batteries referred to above. Deformation of the hair-bearing surface membrane in one direction brings about a reduction in its resistance and a consequent increase in permeability to potassium ions, and hence a depolarisation of the hair cell. In the alternative explanation, an increase in permeability occurs to ions *across the receptor cell membrane lying below the reticular lamina and in contact with cortilymph* (i.e. decrease in R_S in Fig. 14.10). In this case the ion species could be calcium. Both models predict reduction of membrane resistance with depolarising receptor potentials.

As in the case of other receptor cells, the depolarisation of membrane potential releases a chemical transmitter, in the mammalian cochlea probably an amino acid such as aspartic acid or glutamic acid, and this in turn generates characteristically irregular excitatory postsynaptic potentials (EPSPs) in the terminals of cochlear nerve fibres with a synaptic delay of about 0.5 ms. EPSPs larger than a threshold value evoke propagating action potentials, possibly at the point of myelination of the inner radial fibres (beneath the habenular openings, Figs. 14.6, 14.13). The average timing of the EPSPs and action potentials reflects periodicities in the stimulus waveform for frequencies up to about 5 kHz (see p. 284).

The large *efferent* synapses on the base of the outer hair cells liberate acetylcholine as transmitter, which probably opens an ionic channel (R_{EFF} in Fig. 14.10) to chloride ions, thus producing a *hyper*polarisation of the hair cell.

Extracellularly recorded cochlear potentials. It will be evident from Fig. 14.10 that the conductance changes (in either R_A or R_S) responsible for the receptor potentials will generate small time-varying

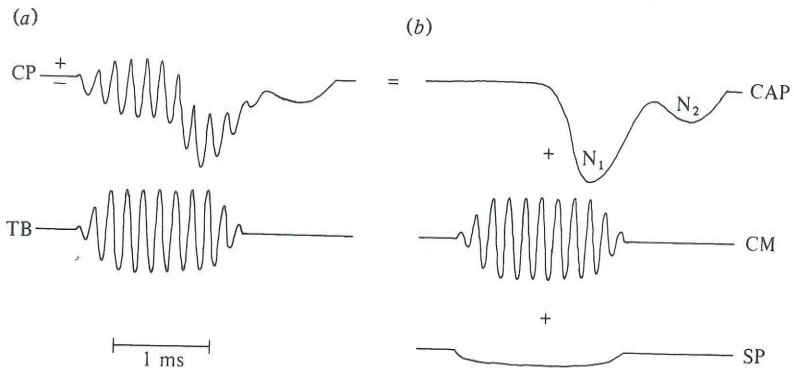


Fig. 14.12. Gross cochlear potentials recorded by an electrode on the cochlea. (*a*) Waveform of cochlear potential (CP) evoked by a short tone burst (TB). (*b*) This waveform can be analysed into three components: gross cochlear action potential (CAP) – a neural response; the a.c. cochlear microphonic (CM) – a receptor response; the d.c. summating potential (SP), probably also a receptor response.

potentials in the scala media, scala tympani and adjacent tissues. These can be easily recorded, together with potentials arising from synchronised action potentials in the cochlear nerve. These cochlear potentials are (Fig. 14.12): the *cochlear microphonic*, *summating potential*, and *gross cochlear action potential*. The cochlear potentials, being more accessible, have been more extensively studied than the intracellular potentials. They can be recorded by an electrode in the cochlear fluids (Fig. 14.10), or on the external wall of the cochlea, and (at very low amplitude) even in the external auditory meatus, and are valuable clinically as the *electrocochleogram*.

The *cochlear microphonic* follows the sound waveform with virtually no latency (the polarity depending upon the electrode location), and is likely to represent, for a single electrode near the round window, the sum of hair cell currents from a large portion of the cochlea, predominantly of the basal turn. The *summating potential*, on the other hand, is likely to reflect the asymmetry of the intracellular receptor potential. There is good evidence to suggest that the cochlear microphonic potential is predominantly generated by the *outer* hair cells.

Because the *gross cochlear action potential* represents the sum of *synchronised* individual action currents of a large number of cochlea fibres, it is observed only at the onset of acoustic stimuli (particularly

to transients such as clicks) subject to a delay of 0.5–1 ms, representing synaptic and conduction delay.

In recordings direct from the inside or from the surface of the cochlea in animals, the cochlear potentials are relatively large (maximum of about 1 mV for cochlear microphonic and cochlear action potential respectively). To obtain the electrocochleogram in patients, however, averaging is required to extract the small cochlear potentials (1–30 μV) from background electrical noise. This entails the use of an *averaging computer* to sum the responses to many repetitions of the signal: the cochlear potentials, being time-locked to the stimulus, sum, whereas the background electrical noise tends to cancel.

Innervation of hair cells

Afferent. The great majority of the 30 000–50 000 afferent fibres in the cochlear nerve arise from the inner hair cells (95% in cat, 85–90% in guinea-pig); see Fig. 14.13. Each of about twenty synapses at the base of each inner hair cell gives rise to an *inner radial fibre*, which immediately leaves the organ of Corti (i.e. radially) via an adjacent or neighbouring opening in the *habenular perforata* into the spiral lamina. Each fibre thus innervates one (in some cases in the guinea-pig two to three) inner hair cell only.

In contrast, the much more numerous outer hair cells are innervated by only 5–10% of the cochlear fibres, which run spirally on their central course (Fig. 14.13). These *outer spiral fibres* innervate about ten outer hair cells in the first 100 μm or so of their spiral course (Fig. 14.13). They then run in the same direction centrally (always apically), without synapsing for considerable distances (0.6 mm in the basal turn of the cat) down the supporting (Deiter's) cells before they cross the floor of the Tunnel of Corti to penetrate a habenular opening, as indicated by the thick lines in Fig. 14.13. Each outer hair cell makes contact with the terminals of several outer spiral fibres.

The afferent fibres from the inner hair cells become myelinated below the habenular openings and pass to their bipolar ganglion cells situated in the spiral *Rosenthal's canal* in the modiolus (see Fig. 14.6). Their central (myelinated) axons leave the cochlea via the modiolus and internal auditory meatus (Fig. 14.3). The fibres from the outer hair cells remain unmyelinated. There is some doubt whether they have functionally significant axons central to their ganglion cell bodies, and what function they serve.

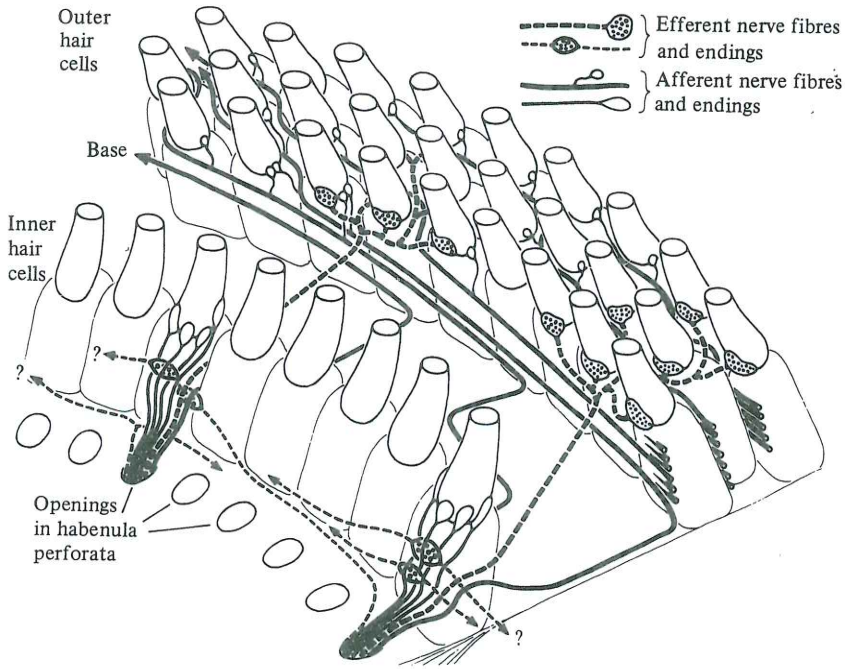


Fig. 14.13. Innervation of inner and outer hair cells. Diagram of surface view of hair cells. For clarity, only a few inner radial fibres ending on the inner hair cells are shown, and a few efferent endings on the outer hair cells. The full spiral extent of the outer spiral fibres cannot be shown on this scale. Note afferent inner radial fibres (shown as thin lines) ending on inner hair cells, with efferent synapses terminating on the afferents themselves; afferent outer spiral fibres (thick lines) crossing the floor of the tunnel of Corti and ascending the supporting cells, with the efferent synapses ending on the outer hair cell. (From Spoendlin in *Frequency Analysis and Periodicity Detection in Hearing*. p. 2. Sijthoff, 1970.)

Efferent. A very substantial proportion of the synaptic terminals in the cochlea, particularly on the outer hair cells, are *efferent*, that is, are part of the descending, centrifugal, auditory pathways (Fig. 14.1). The *olivocochlear bundle*, as its name implies, arises in the superior olive region of the brainstem as *uncrossed* (ipsilateral) and *crossed* (contralateral) bundles and enters the cochlea via the vestibular nerve. The crossed bundle runs predominantly but not exclusively to the outer hair cells (Fig. 14.6) ending in huge terminals (Fig. 14.13), dwarfing and outnumbering those of the afferents. The remainder of

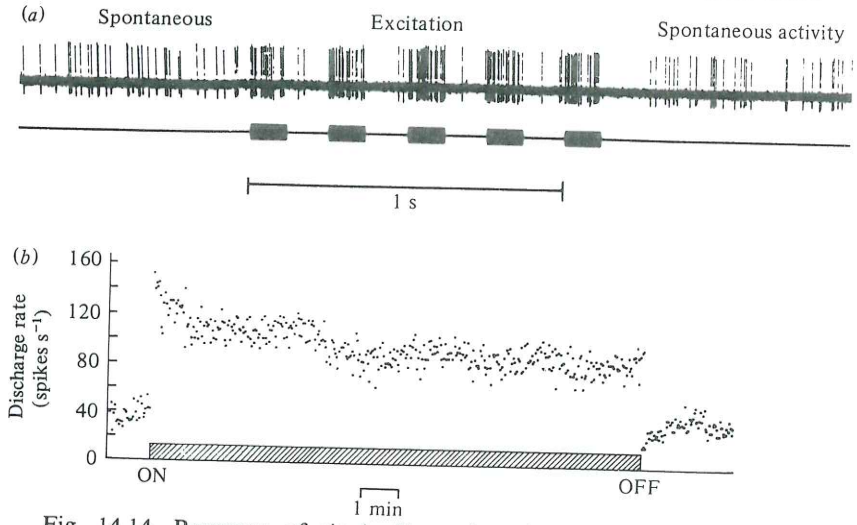


Fig. 14.14. Response of single fibre of cochlear nerve to sound. (*a*) Spontaneously active fibre responds to 5 tone bursts by producing bursts of action potential spikes. Note reduction in activity (off-suppression) following each burst of activity. (Guinea-pig: From Evans (1975). Cochlear nerve and cochlear nucleus. In *Handbook of Sensory Physiology*. Vol. V/2. Chapter 1. Springer-Verlag.) (*b*) Time course of response to a continuous tone of 13-min duration. Each point indicates the number of spikes counted in 1 s. Note discharge rate is maximum at onset of stimulus, and progressively reduces thereafter (adaptation); transient reduction in probability of discharge on terminating stimulus (off suppression). (Cat: After Kiang *et al.* (1965). *Discharge patterns of single fibres in the cat's auditory nerve*. M.I.T. Press.)

the efferent fibres form the *inner spiral bundle* ending on the inner hair cell afferent fibres themselves (Fig. 14.13).

14.4. COCHLEAR NERVE

There is now available a very large body of quantitative data on the responses of individual cochlear nerve fibres (mainly in cat, guinea-pig and squirrel monkey) to a wide variety of stimuli. In the absence of sound stimulation, the majority of cochlear fibres are spontaneously active (*spontaneous activity*, Fig. 14.14). To tones of appropriate frequencies and intensity, all fibres give a characteristic response: *excitation* lasting the duration of the stimulus, after a latency of 1–10 ms (Fig. 14.14*a*). Both the spontaneous and the evoked discharges are characteristically irregular.

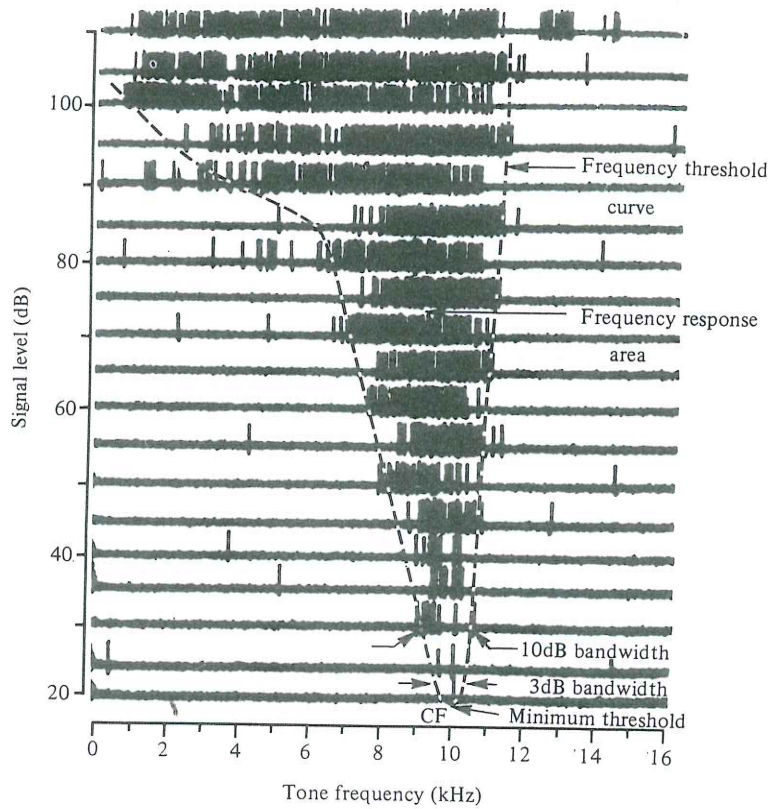


Fig. 14.15. Frequency response area of single fibre in cochlear nerve. Frequency response area is mapped out by sweeping a continuous tone successively upward and downward in frequency, increasing the tone level by 5 dB at the end of each sweep. At the lowest intensities, the fibre responds to a narrow band of frequencies at 10 kHz. This is its *characteristic frequency* (CF). At higher levels, the excitatory frequency band becomes progressively wider. For fibres of this high characteristic frequency, the response area is asymmetrical, extending on the low frequency side to form a low frequency 'tail' above about 80 dB SPL. The boundary of the frequency response area is called the frequency threshold curve (FTC) or 'tuning curve'. (Minor deviations of the responses from the curves are caused by 'off-suppression' effects working in opposite directions in each sweep.) Note linear frequency scale. (Guinea-pig: from Evans (1972) *Journal of Physiology*, *226*, 263.)

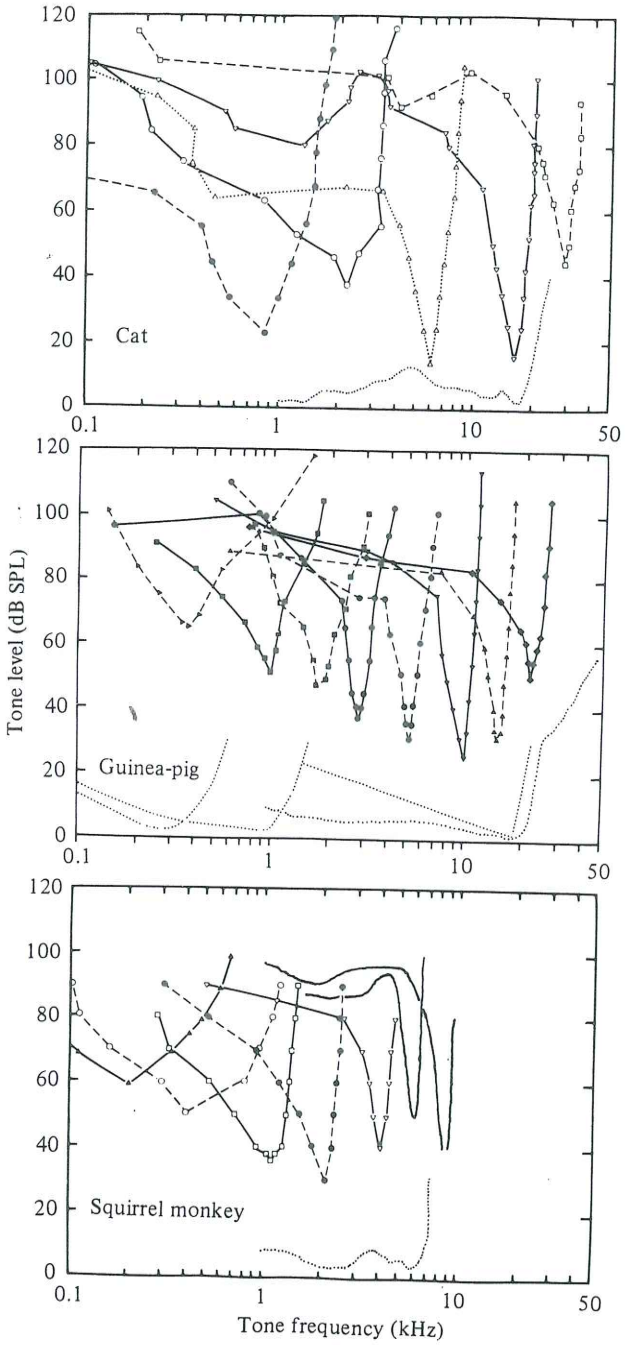
In common with all receptor neurons, the rate of discharge evoked by a continuous stimulus *adapts* in time (Fig. 14.14*b*). The decrement in rate is approximately related to the log of time. On termination of the stimulus, a brief period of depression of spontaneous activity (and excitability) ensues.

Frequency selectivity

The range of frequencies that will evoke responses from any given cochlear fibre is relatively limited. Fig. 14.15 shows the responses of a single fibre innervating the 10 kHz place in the organ of Corti, to sweeps in continuous tone frequency, in each direction. At the lowest signal level illustrated (+20 dB) the fibre responds, at its *minimum threshold*, to 10 kHz. This frequency of maximum sensitivity is termed the fibre's *characteristic frequency* (CF). At progressively higher stimulus levels, the effective frequency band grows wider, so that at the highest level tested, 90 dB above minimum threshold, frequencies from 1 to 11 kHz evoke responses. The range of effective frequencies and intensities map the *frequency response area* of the fibre (equivalent to the *receptive field*) and the threshold boundary, the *frequency threshold* ('tuning') *curve* (FTC). Fig. 14.16 depicts families of FTCs from cochlear fibres originating from different points along the cochlea, from guinea-pig, cat and monkey. For those fibres with characteristic frequencies above 2 kHz, the FTCs are asymmetrical, and represent quite remarkable filters. They have very steep cut-off slopes on the high-frequency side, and low-frequency cut-offs that are somewhat less steep at lower sound levels, but which suddenly decrease in slope, above 70–100 dB SPL, thus forming a low-frequency 'tail' to the response area. At least for 40–60 dB or so above minimum threshold, the cochlear fibres act as narrow band filters, having bandwidths ranging from about $\frac{1}{2}$ octave at 0.2 kHz to $\frac{1}{10}$ octave at 20 kHz. Their shapes match those of the isoresponse curves of the receptor potentials of inner hair cells (Fig. 14.11), and as we shall see, the analogous psychoacoustic 'tuning curves' (Fig. 15.1).

In contrast, the tuning of the cochlear fibres is sharper than existing measurements of the amplitude response of the basilar membrane (dotted lines in the lower part of each panel of Fig. 14.16).* At least part of this difference in tuning may be due to the action of a metabolically active, so-called 'second filter' mechanism

* Recent studies suggest that the basilar membrane tuning is much sharper than that indicated in Figs. 14.9 and 14.16, and that active mechanisms are involved. For references, see p. 306.



in the cochlea (the first filter being that of the basilar membrane). Some candidates proposed for this frequency sharpening mechanism are the 'micromechanics' of the hair-cell: tectorial membrane region (p. 267), electrical tuning of the hair cells (p. 270), and even an active motion of the hair cells or the cilia themselves.

The minimum threshold (i.e. at the FTC tip) of cochlear fibres follow closely the behavioural audiogram of the species (Fig. 14.5*b*). As indicated earlier (p. 257), this overall frequency sensitivity is determined largely by the characteristics of the outer and middle ears. Within an individual ear, the range of minimum thresholds at any frequency is relatively restricted, to 20 dB or so. Interestingly, the most sensitive fibres tend to have higher rates of spontaneous activity than the less sensitive.

The representation of frequency along the cochlear partition and the discrete innervation patterns of the inner radial fibres imply that the characteristic frequency of a cochlear fibre depends upon where it originates. Fibres innervating the apical end have the lowest characteristic frequencies (in the cat about 0.1 kHz) and those from the basal end, the highest (in the cat about 50 kHz). Because this mutual relationship of cochlear fibres is preserved in the cochlear nerve trunk (in the manner of a spiral) and in its termination in the various subdivisions of the cochlear nucleus (§14.5), and because cochlear fibres are relatively narrowly tuned, the peripheral auditory system is *cochleotopically* and *tonotopically* organised, i.e. stimuli of different frequencies (*tonos*) evoke activity at systematically ordered positions (*topos*) along the neural array. Thus, in principle, the frequency of stimulus is coded by the *place* of neural activity (p. 315).

Fig. 14.16. Families of frequency threshold curves for single cochlear nerve fibres in cat, guinea-pig and squirrel monkey. Each curve represents the frequency threshold curve of a different fibre, the fibres chosen to cover a wide range of characteristic frequencies, on a logarithmic frequency scale. Below the neural curves are arbitrarily positioned analogous response curves (dotted lines) for the basilar membrane, to show the difference in tuning between basilar membrane (low-pass) and the cochlear nerve fibres (narrow-band). Neural data: cat, Kiang *et al.* (1967) *Journal of the Acoustical Society of A.nerica*, *42*, 1341; guinea-pig, Evans (1972) *Journal of Physiology*, *226*, 263; squirrel monkey, Rose *et al.* (1971) *Journal of Neurophysiology*, *34*, 685; Geisler *et al.* (1974) *Journal of Neurophysiology*, *37*, 1156. Basilar membrane data: cat, Evans & Wilson (1975) *Science*, *190*, 1218; guinea-pig, Wilson & Johnstone (1972). *Hearing Theory*. IPO Eindhoven; squirrel monkey – Geisler *et al.* (1974) *Journal of Neurophysiology*, *37*, 1156.

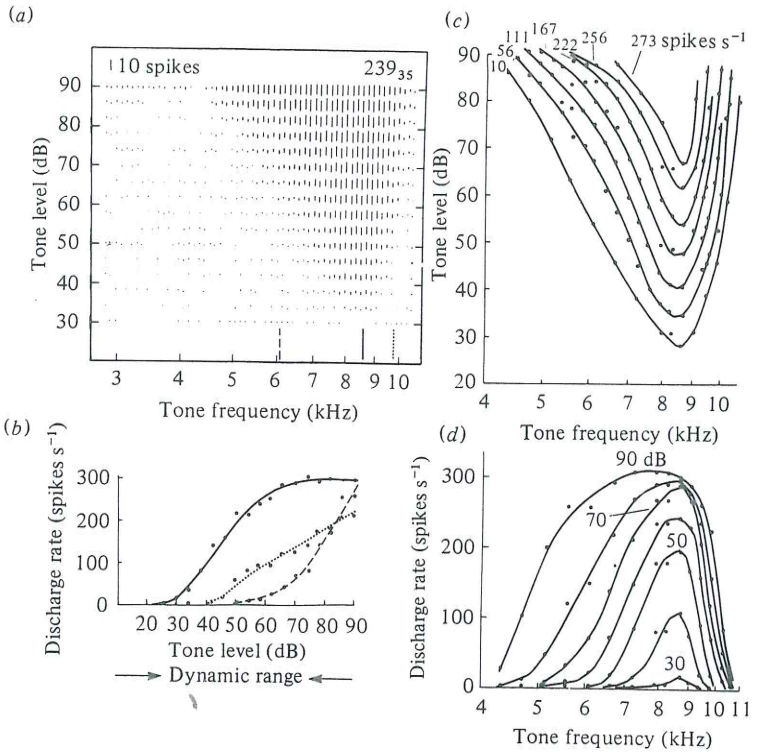


Fig. 14.17. Response of single cochlear nerve fibre as function of frequency and intensity. (*a*) Frequency response 'map' of fibre. The length of each vertical line indicates the average number of spikes evoked by a 50-ms duration stimulus at the frequency and intensity indicated by the centre of the line. (*b*) Vertical 'sections' through *a* at frequencies indicated by dashed, continuous and dotted lines respectively. These are known as rate–intensity or rate–level functions. Note restricted *dynamic range* over which the fibre can signal intensity at its characteristic frequency, and *saturation* of discharge rate at high levels. (*c*) Isorate contours. Each contour indicates the tone frequencies and intensities evoking a given discharge rate. These are obtained by taking horizontal 'sections' through a family of rate-level functions as in (*b*). (*d*) Iso-intensity contours, i.e. horizontal 'sections' through (*a*). Note flattening of the contours at the highest levels, because of saturation of the response, and a shift in frequency of maximum response. (Cat: from Evans (1978) *Audiology*, *17*, 369.)

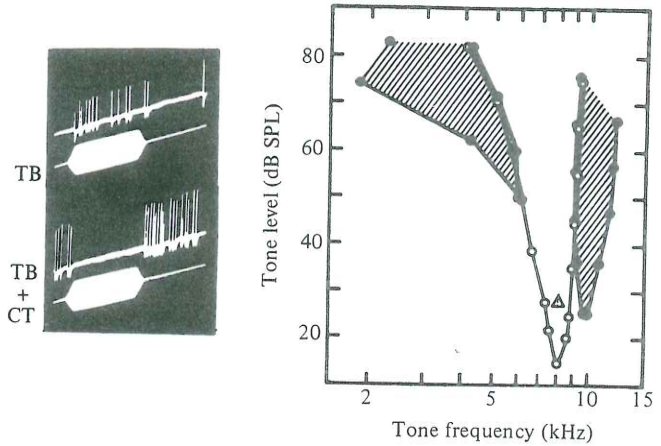


Fig. 14.18. Lateral suppression of cochlear fibre response. The response to stimulus tone (indicated by the triangle at 8 kHz, the fibre's characteristic frequency) can be suppressed by a second stimulus falling within the shaded areas. The suppression areas slightly overlap the frequency threshold curve (unshaded area) so that a tone burst that excites the fibre when presented alone (TB) suppresses the response to a continuous exciting tone (CT). (From Evans (1975). *Cochlear nerve and cochlear nucleus*, in *Handbook of Sensory Physiology*. Vol. V/2 Chapter 1, Springer-Verlag; After Nomoto *et al.* (1964). *Journal of Neurophysiology*, *27*, 768; Arthur *et al.* (1971) *Journal of Physiology*, *212*, 593.)

The above description of cochlear fibre tuning refers to *threshold* characteristics. The suprathreshold response is also tuned, up to a point (Fig. 14.17*c*, *d*). Most fibres have a limited dynamic range over which they can signal the level of a tone stimulus of given frequency. Thus, at the characteristic frequency of the fibre illustrated in Fig. 14.17, stimulus levels between about 25 and 70 dB SPL evoke a monotonically increasing discharge rate (Fig. 14.17*b*). Below this range, the stimuli are subthreshold, and above, the unit's discharge rate *saturates*, owing to some inherent non-linearity in the transducer and/or synaptic mechanisms. This limited dynamic range makes it difficult to understand how the intensity of signal components is peripherally coded; this problem is discussed in Chapter 15.

Another instance of cochlear non-linearity is the phenomenon of *two-tone or lateral suppression* (Fig. 14.18). Here, the response of a cochlear fibre to a tone (at the triangle in Fig. 14.18) can be suppressed by a tone falling within frequency regions flanking the

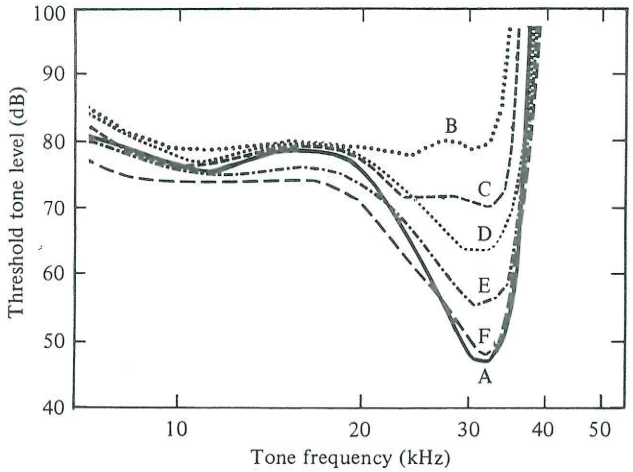


Fig. 14.19. Physiological vulnerability of cochlear tuning. Effects obtained on the tuning of a single cochlear nerve fibre in cat, from an intra-arterial injection of furosemide, an ototoxic diuretic known to cause reversible hearing loss in man. A (continuous line) is the fibre's frequency–threshold curve before injection, B and C are the curves obtained 1 and 2 min respectively after injection. Note the progressive loss of the low threshold, sharply tuned 'tip' segment of the frequency threshold curve until, at C, only a high threshold, broad curve remains. Curves D, E and F indicate that the effects on the neural tuning are reversible at this dosage. They were obtained 5, 7 and 20 min after the injection. (From Evans & Klinke (1974) *Journal of Physiology*, *242*, 129P.)

response area, as indicated by the hatched areas in Fig. 14.18. Because these suppressive areas slightly overlap the frequency threshold curve, the suppressive stimulus itself can actually excite the fibre in the absence of the CF tone (Fig. 14.18). The mechanism underlying this phenomenon is not clearly understood.

Physiological vulnerability of cochlear frequency selectivity

The cochlear fibre responses (and hair-cell receptor potentials) are sharply tuned and highly sensitive only if the cochlea is in normal physiological condition. Deterioration in the condition of the cochlea can produce selective loss of the low-threshold, sharply tuned segment of the FTC (Figs. 14.19, 14.20). This loss may be acute, i.e. short-lived and reversible, such as when it results from brief hypoxia, or from systemic administration of doses of ototoxic agents such as

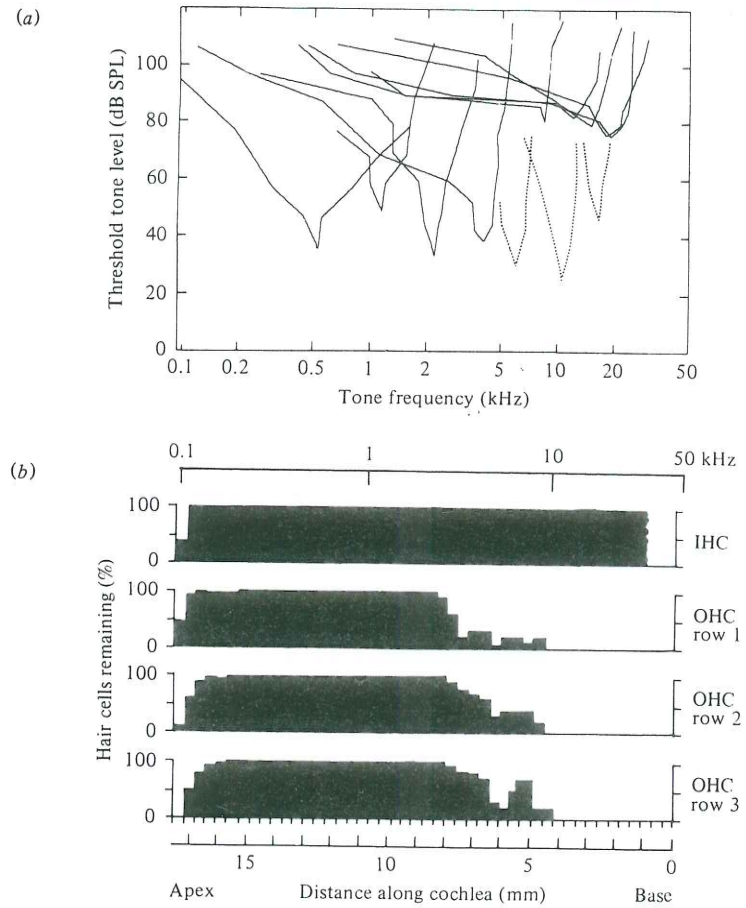


Fig. 14.20. Effects of hair-cell damage on cochlear fibre tuning. (*a*) Continuous curves are the frequency–threshold curves obtained from cochlear nerve fibres in a guinea-pig ear damaged by long-term injections of kanamycin, an antibiotic of the ototoxic streptomycin family. (*b*) The black areas below indicate the proportion of inner and outer hair cells remaining in the cochlea at the time of the physiological recording. All of the outer hair cells are missing from the basal half turn (first 4–5 mm) of the cochlea. The frequency–threshold curves of fibres with characteristic frequencies corresponding to the damaged regions have high thresholds and are broad; the dotted outlines indicate the appearance of normal curves. (From Evans & Harrison (1976) *Journal of Physiology*, *256*, 43P.)

furosemide (a powerful diuretic used clinically in renal failure, Fig. 14.19). On the other hand, permanent loss of tuning can be caused by local damage (e.g. haemorrhage), or by the systemic administration of the aminoglycoside antibiotics (streptomycin, neomycin, gentamycin, kanamycin) as shown in Fig. 14.20*a*. Morphological evidence of damage to the organ of Corti often occurs in these cases, particularly after long-term administration of agents like kanamycin (Fig. 14.20*b*). The basal turn and especially the outer hair cells are generally affected first as shown by the *cochleogram* of Fig. 14.20*b* indicating the percentage of hair cells remaining. Curiously, under these conditions, the threshold and tuning properties of the cochlear fibres recorded from the same cochlea correlate more with the presence and absence of the *outer* hair cells than with the *inner* hair cells that the fibres innervate (assuming that the fibres recorded from originate from the inner hair cells, and that the latter are *functionally* normal). This suggests that the normal sharply tuned properties of the inner hair cells depend in some way upon the integrity of the outers. This could afford the outer hair cells with an important modulating function, in view of their apparently insignificant afferent, but powerful efferent innervation.

This *physiological vulnerability* of the cochlear filter is of considerable significance for our understanding of the nature of sensorineural hearing loss of cochlear origin (p. 311).

Temporal patterning of discharges: coding of stimulus period

So far, we have limited discussion of cochlear-fibre responses to the *mean rate* of the discharges. For suprathreshold stimuli of frequency below about 5 kHz, impulses in cochlear fibres show a temporal pattern (Fig. 14.21); there is a preference for the fibre to discharge in a given half cycle of the stimulus period (Fig. 14.21*a*). This is shown more clearly in the *period histogram* of Fig. 14.21*b*, where the number of discharges per unit of time across the cycle of the stimulus is accumulated, and plotted as a function of the stimulus phase. The phenomenon is known as *phase locking* of the discharges. Thus, the probability of discharge is roughly proportional to the half-wave rectified stimulus waveform (Fig. 14.21*b*). This 'phase-locking' can be demonstrated by another type of histogram (Fig. 14.21*c*) where the number of *intervals* of different durations between action potentials are accumulated – the *interspike interval histogram*. This shows an exponential relationship between the number of intervals and their

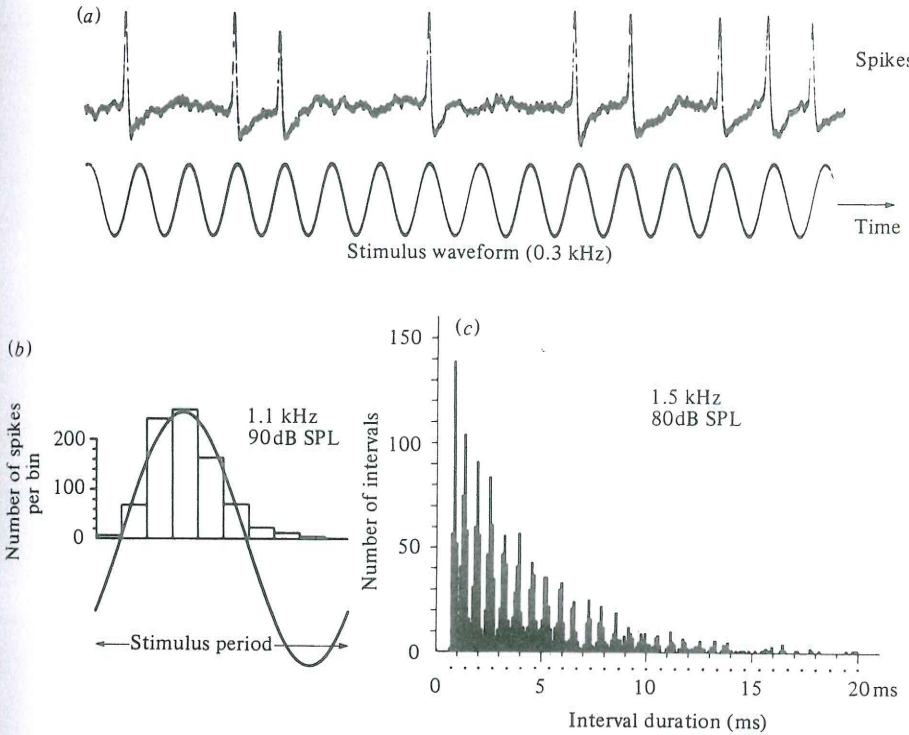


Fig. 14.21. Temporal patterning of discharges of single cochlear nerve fibres. (*a*) Response in time of a single cochlear fibre to a low-frequency continuous tone. Note that the spikes are 'phase-locked' to a given half-cycle of the stimulus waveform. (Guinea-pig: from Evans (1975). *Cochlear nerve and cochlear nucleus* in *Handbook of Sensory Physiology*. Vol. V/2 Chapter 1, Springer-Verlag.) (*b*) 'Period histogram' of responses similar to those in (*a*). The height of each bin indicates the number of spikes falling in that portion of the period of the waveform. Note that the histogram approximately follows a half-wave rectified version of the stimulus waveform (which has been phase-shifted to match the histogram). (Squirrel monkey: Rose *et al.* (1971). *Journal of Neurophysiology*, *34*, 685.) (*c*) Interspike interval histogram indicating number of intervals of given duration obtained from response of a cochlear fibre to continuous tone of 1.5 kHz. The separation between the histogram peaks reflects the period of the stimulus (indicated by the dots under the abscissa). (Squirrel monkey: from Rose *et al.* (1968) in *Hearing Mechanisms in Vertebrates* p. 144. Churchill.)

duration,* modulated so that the intervals chiefly represented are those corresponding to the *period* of the stimulus and its multiples.

Interestingly, some 'phase-locking' occurs in fibres at stimulus levels as much as 20 dB *below* the threshold for a change in the mean discharge rate. Phase-locking also persists at sound levels *above those causing saturation of the mean discharge rate* (although the *degree* of phase-locking saturates at lower levels).

Complex waveforms comprising several frequencies below 5 kHz, produce responses where the waveforms, after modification by the cochlear filter, are represented in the period histogram and in the distribution of interspike intervals. This representation persists relatively unaffected at sound levels above saturation of the mean discharge rate, and is therefore remarkably robust. This temporal information is certainly important for the localisation of the sources of a sound (§15.4) but whether and how it is used in analysing frequency is not known.

Response to click stimuli

Click stimuli, generated by brief pulses, are in principle wide-band. They can therefore excite every cochlear fibre, providing the energy filtered by the cochlear filter exceeds the threshold of that fibre. As expected, for fibres of low CF, the 'ringing' of the cochlear filter is manifested in the temporal response patterns of the discharges as a periodic enhancement of discharge probability (Fig. 14.22*a*, *b*). For fibres with CF higher than about 4–5 kHz (the approximate limit of 'phase-locking'), this periodicity is blurred out.

Summary

This brief account of the response properties of cochlear fibres has emphasised a number of features that make the task of predicting their responses to complex sound more straightforward. *To a first approximation*, these responses can be predicted on the following basis: (i) *linear filtering* with a filter shape corresponding to the frequency threshold curve; (ii) half-wave *rectification* of the waveform; (iii) a *logarithmic transform* to represent the approximately linear

* This exponential relation is expected for a Poisson process in which the occurrence of an impulse is equally likely at any time, the *probability* of occurrence being the only quantity that can be specified. For cochlear nerve fibres this probability is primarily determined by the frequency and intensity of the sound, but is also modulated by its phase as shown in Fig. 14.21. Note that there is a short period following an impulse within which the probability of a second impulse is reduced by refractoriness.

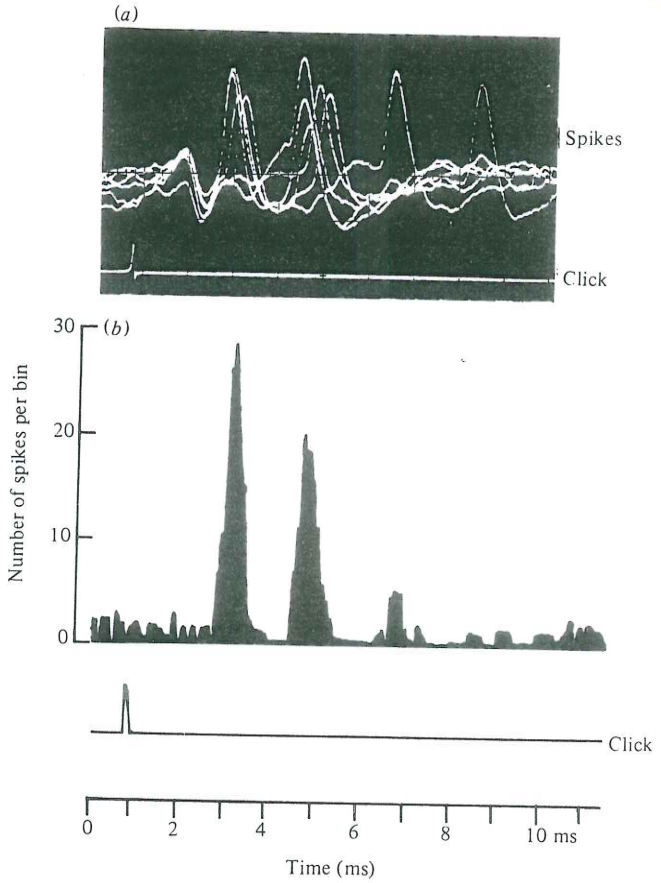
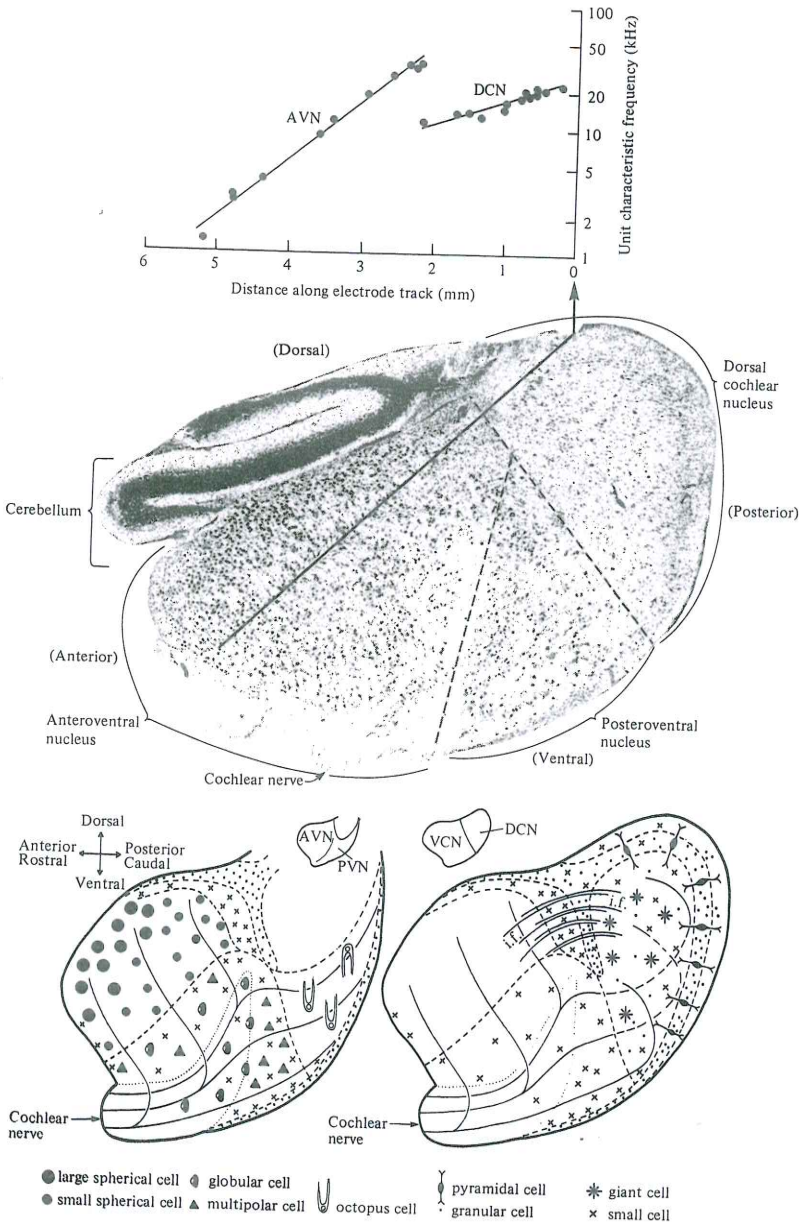


Fig. 14.22. Response of single fibres in cochlear nerve to click stimuli. (*a*) Superimposed records of response of low-frequency fibre (characteristic frequency, 0.5 kHz) to 6 click stimuli. Note grouping of discharges at preferred times following the click. This indicates the 'ringing' properties of the cochlear filter. (*b*) Post-stimulus time histogram of many responses (as in *a*). Height of each bin indicates the number of spikes elicited at the given time interval following the click stimulus (at 1 ms). Note periodicity of discharge probability having a period (2 ms) that corresponds to the reciprocal of the fibre's characteristic frequency (0.5 kHz). Thus the 'ringing' frequency and characteristic frequency both reflect the same cochlear filter.



relation between discharge rate and decibel stimulus level within the limits of threshold and saturation; (iv) a low-pass (smoothing) filter which limits the frequency range over which 'phase-locking' occurs; (v) a *probabilistic spike generator*, in which the probability of discharge is a function of the filtered, rectified and transformed waveform.

The cochlear nerve can therefore be considered to represent a *filter bank*, each fibre representing a narrow band-pass filter covering overlapping but slightly different frequencies from its neighbour. Both the mean discharge rate (below saturation) and, for stimuli with frequency components below about 5 kHz, the temporal discharge patterns of each fibre, are determined by the relative magnitude of the frequency components passed by its filter.

14.5 COCHLEAR NUCLEUS

Organisation

In the cochlear nucleus we see the first signs of divergence and parallel specialisation of functions in the auditory system (Figs. 14.1, 14.2). The cochlear nerve bifurcates on entry to the nucleus, sending terminals in *parallel* to the three main subdivisions of the nucleus (Fig. 14.23): the antero- and postero-ventral divisions, and the dorsal division. Because of the orderly arrangement of the cochlear fibres and of their branching, each subdivision is tightly cochleotopically organised (Fig. 14.23): i.e. the cells of each subdivision have characteristic frequencies that are logarithmically related to linear distance

Fig. 14.23. Anatomical organisation of cochlear nucleus. The cochlear nerve enters the nucleus anteriorly and ventrally. Centre of figure: sagittal section though cat cochlear nucleus stained to show cell bodies (Nissl stain). Dashed lines indicate major subdivisions: ventral cochlear nucleus (VCN), subdivided into anteroventral (AVN) and posteroventral (PVN) nuclei, and dorsal cochlear nucleus (DCN). Note the lamination of the dorsal nucleus below the surface, and the separation of the more homogeneous anterior and posterior divisions of ventral nucleus (by root of cochlear nerve). Lower third: major classes of cell types and their locations, according to Osen (1970) *Italian Archives of Biology*, *108*, 21). i.f., intrinsic fibres interconnecting anteroventral and dorsal nuclei. Upper third: plot of characteristic frequencies of cochlear nucleus cells at the position indicated along the electrode track (from dorsal to ventral) shown as continuous straight line in the centre figure. Note tight cochleotopic organisation of each major subdivision.

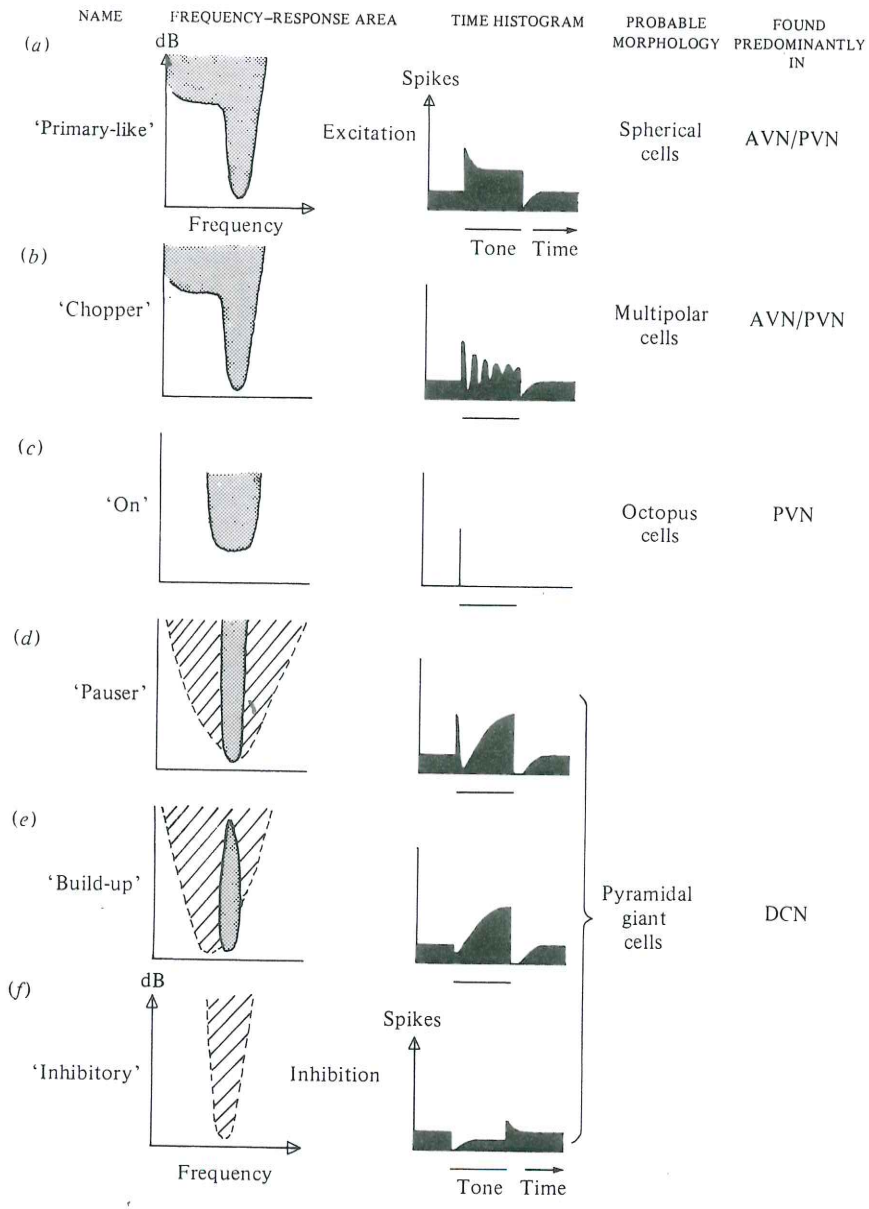


Fig. 14.24. Diagram of frequency and time responses of major types of cells in cochlear nucleus. Left-hand column: diagrams of frequency response areas (axes as in *f*) indicating excitatory (stippled) and inhibitory (shaded) frequency–response areas of each type of cell. Inhibition alone (*f*), or combined with excitation (*d*, *e*), is found predominantly in the morphologically more complex dorsal nucleus. Middle column: diagrams of peristimulus-time histograms of response to characteristic frequency tone of each type of cell (axes as in *f*). The shape of the PST histogram gives rise to the 'nick-name' of each cell type. Right-hand column: probably cell types and location.

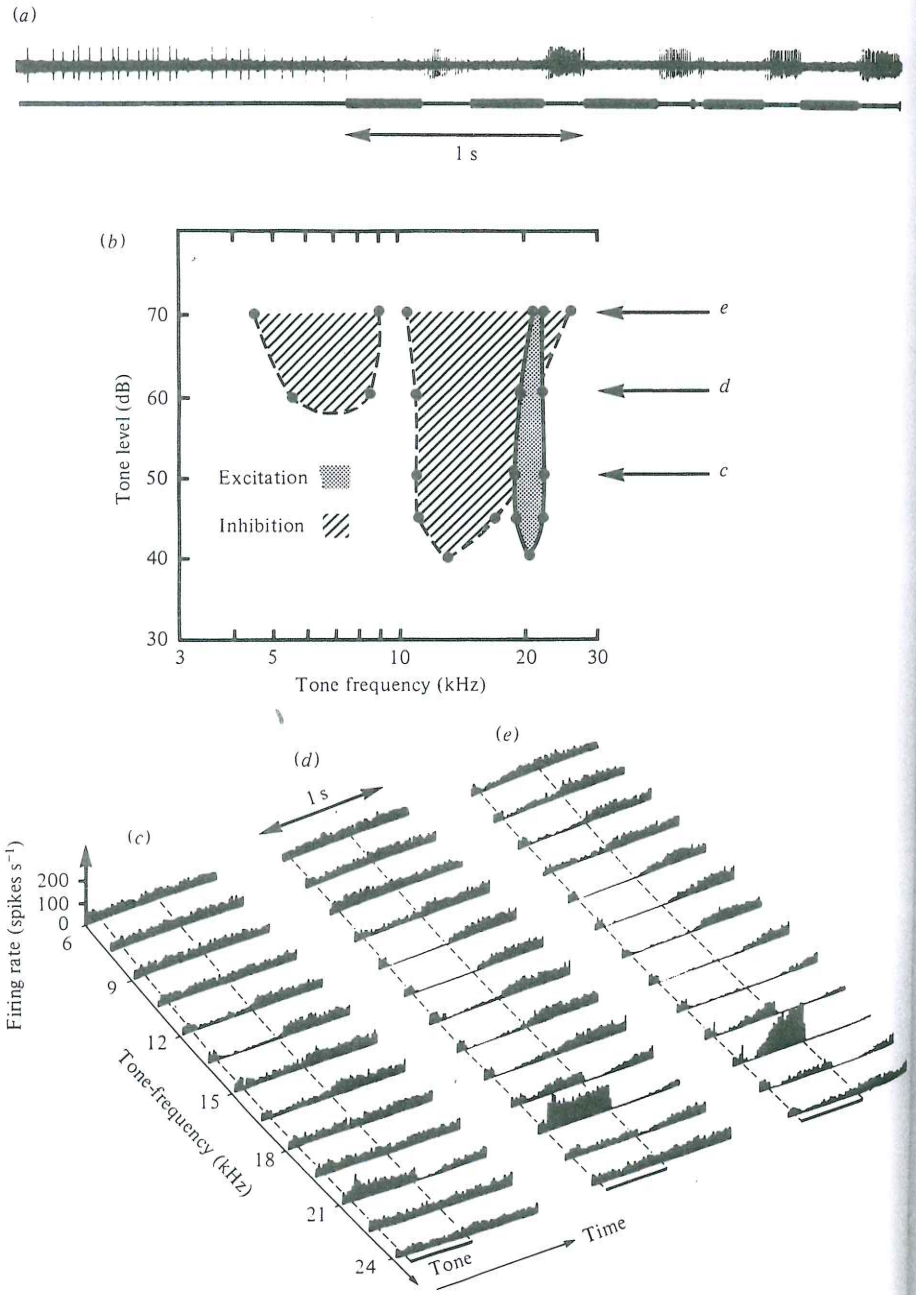
in the nucleus. (In Fig. 14.23 the electrode track passes only through the dorsal and anteroventral divisions.)

Each subdivision is characterised by a particular distribution of the different cell types (Fig. 14.23). Thus the ventral divisions are the more homogeneous, and the dorsal is the more complex, being laminated. Fig. 14.23 also illustrates the strong bundle of intrinsic fibres (i.f.) connecting the dorsal and ventral divisions.

Subdivision of function in the cochlear nucleus

As might be expected, the different cell morphologies reflect different synaptic dispositions and therefore response properties. These in turn presumably imply specialisations of function (see below). The main correlations can be summarised as follows (Fig. 14.24). The majority of the cells in the ventral nucleus have discharge properties not very different from those of the cochlear fibres. They are hence termed 'primary-like'. Exceptions are the 'chopper' and 'on' cells. The latter are the only type of response found in the 'octopus cell' region. These cells, because of their large dendritic extensions cutting across a relatively large extent of the tonotopically organised afferent array, are also unusual in that they have very broad frequency response areas, and can follow the individual pulses or cycles of click or tone stimuli at repetition rate of up to 800 s^{-1}.

In the dorsal nucleus, more complex responses of longer latency are found and inhibitory inputs are particularly apparent. These inhibitory inputs are such that a *single* tone can reduce – often completely – the cell's activity, as in Fig. 14.25. In some cases, the termination of the inhibition is accompanied by an 'off' response as in Fig. 14.25*a*. It is thus different from the 'two-tone' suppression at the cochlear nerve level, and is likely to result from lateral inhibitory synaptic mechanisms, analogous to those in the retina (Chapter 6). The intrinsic fibre projection (i.f. in Fig. 14.23) from the ventral nucleus represents a delayed inhibitory input to the dorsal nucleus cells.



In the frequency response areas, the extent of inhibition ranges from one or two relatively restricted 'sidebands' contiguous with an excitatory response area ('pauser', Fig. 14.24), through extensive sidebands restricting the width of the excitatory area at high levels (Fig. 14.25), to a solely inhibitory response area ('inhibition', Fig. 14.24). This separate mapping of excitatory and inhibitory response areas does not imply that they are independent; on the contrary, the 'excitatory' responses are commonly a complex interaction of excitation and inhibition (as in the time histograms of Fig. 14.25). In addition, the inhibition commonly produces a non-monotonic relation between discharge rate and intensity, in contrast to the situation at the cochlear nerve: in other words, the discharge rate at high intensities may be lower than at intermediate intensities.

This lateral inhibition may enhance temporal contrast, and contrast across the frequency spectrum; it may increase the dynamic range at this level (p. 323), and increase sensitivity to frequency *change* in one direction, depending upon the asymmetry of inhibition. (This is presumably analogous to the enhancement, by asymmetrical lateral inhibition, of response to one direction of movement in the rabbit retina; see Chapters 1 and 8.)

There are at least three parallel subsystems at work in the cochlear nuclei, each having its own specific output pathway (Figs. 14.1, 14.2). The trapezoid body, forming a *ventral acoustic pathway*, carries 'primary-like' information from the ventral cochlear nucleus to the superior olivary nuclei. Likewise, the *intermediate acoustic stria* conveys the 'on' type response of the octopus region to the superior olivary nucleus. These two pathways specialise in transmitting, with the minimum of delay, information necessary for the correlation of

Fig. 14.25. Complexity of neural inhibition in two neurons from dorsal cochlear nucleus. (*a*) 'Off' response following inhibition of first unit's spontaneous discharge by tone at characteristic frequency. (*b*) Map of frequency–response areas of excitation (stipple) and inhibition (cross-hatch shading) of second neuron. (*c*, *d*, *e*) Array of peristimulus-time histograms showing time course of responses mapped out in *b*. Note complexity of time course of inhibition during and after tones. Note also that 'excitatory' response areas e.g. at 21 kHz, 70 dB SPL (*e*) represent a complex mixture of excitation and inhibition in time. This figure also indicates that a cell's response type depends on frequency and intensity (e.g. 'pauser', 'inhibition', 'build-up'). (Cat: after Evans & Nelson (1973). *Experimental Brain Research*, *17*, 402.)

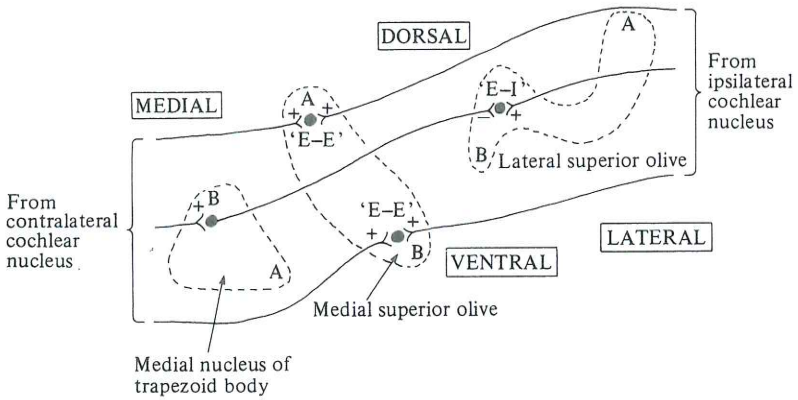


Fig. 14.26. Organisation of Superior Olivary Complex. Right superior olive of cat. Each subdivision is organised cochleotopically, i.e. the characteristic frequencies of the constituent cells are arranged so that in the medial nucleus of the trapezoid body and lateral superior olive, low frequencies (the cochlear apex, A) are represented laterally and the cochlear base (B) medially. The reverse is the case in the medial superior olive. The inputs to the E–E cells and E–I cells are shown.

inputs from the two ears (see next section). They are, presumably, chiefly concerned with the neural analysis of the *position* of stimuli in acoustic space. In contrast, the *dorsal acoustic pathway*, formed by the *dorsal acoustic stria* from the dorsal cochlear nucleus, is presumably chiefly concerned with the analysis of the *pattern* of stimuli. Whether the separate functional identities of these pathways are retained at higher levels of the auditory system is unknown.

14.6. SUPERIOR OLIVARY NUCLEI

This collection of nuclei (Fig. 14.26), lying in the ventral pons, receives input from both ears, and therefore represents the lowest neural level at which correlations are made between signals arriving at the two ears. This appears to be its chief function.

Three modes of binaural interaction are found in this nucleus. The first two are mutually exclusive, involve cells having characteristic frequencies above 1 kHz, and are concerned with correlating *intensity* between the two ears. In the first mode, characteristic of cells having excitatory inputs from both ears (called *E–E cells*) the cells are more sensitive to variations in binaural intensity (i.e. variations that are

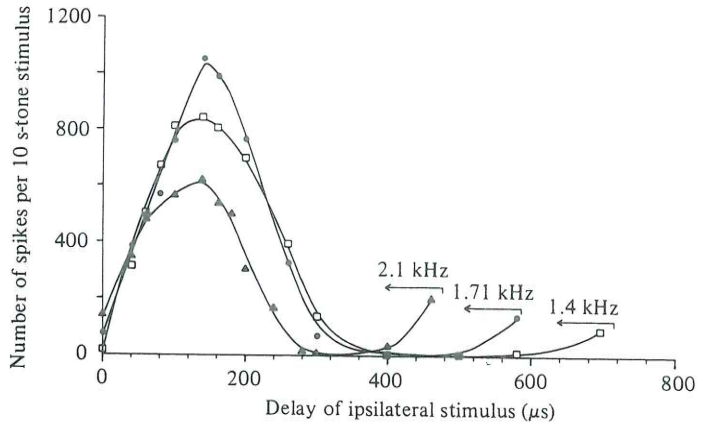


Fig. 14.27. 'Critical delay' cell. Response of a cell in inferior colliculus as a function of the time delay between signals arriving at the two ears. A response maximum occurs at the so-called critical delay of 160 μs largely irrespective of the frequency of the tone (or its level). Similar responses are found in the superior olive. (Cat: After Rose *et al.* (1966). *Journal of Neurophysiology*, *29*, 288.)

correlated at the two ears) than to differences in intensity between the two ears. In the second mode, the converse is the case: the cells are more sensitive to changes in the intensity *difference* between the two ears than to changes in the *average* binaural stimulus level. These are cells receiving excitation from one ear and inhibition from the other (*E–I cells*). This second mode is particularly relevant for the localisation of high-frequency sounds, where head-'shadowing' and the effects of the pinna cause significant differences in interaural intensity, depending upon the location of the source (§15.4).

The third mode of interaction is found in both E–E and E–I type cells having characteristic frequencies below about 1 kHz: sensitivity to the interaural *time delay* between sounds reaching the ears. In the case of the E–I cells, whether the cell is predominantly excited or inhibited depends upon the interaural delay. For some cells, the maximum response occurs at a given interaural delay largely irrespective of the intensities at the ears, and the frequency of the signal. These have been termed 'critical delay' cells (see Fig. 14.27). For others, the interaural time and intensity differences interact in such a way that the effects of one can be traded against the other. Thus, the effects of increasing the intensity at one ear can be offset by delaying the sound to that ear. This '*time-intensity trading*' is of

the same order of magnitude as that required in psychophysical experiments to maintain a sound image in the same subjective location.

Both medial and lateral divisions of the superior olive are tonotopically organised, indicating a precise convergence of input fibres from the cochlear nuclei of the two sides.

14.7. INFERIOR COLLICULUS AND MEDIAL GENICULATE NUCLEUS

The auditory system differs from the visual system particularly at this level, in that the *inferior colliculus* is an obligatory station on the pathway to the cortex (Figs. 14.1, 14.2). The fibres of the lateral lemniscus end predominantly in its central nucleus, sending collaterals to the thin surrounding external (pericentral) nucleus which also receives input from the central nucleus. Both nuclei are cochleotopically organised.

As expected from the convergence of dorsal and ventral pathways and pathways from the superior olives, a great variety of types of response are encountered at this level. Of particular note are neurons that do not respond to steady tones but can be activated by frequency or amplitude modulation. In the bat, neurons have been described that are sensitive even to the *form* of the amplitude modulation. These neurons may subserve the analysis of the rate of rise of stimulus transients. As in the superior olive, the pattern of excitation and/or inhibition in many cells depends upon binaural influences (Fig. 14.27). However collicular neurons, by comparison, demonstrate a higher degree of sensitivity to interaural time or intensity differences. Interestingly, relatively few of the neurons sensitive to interaural time differences show phase-locked responses at low-frequency tones.

Another type of binaural processing has been reported at this level. This is the presence of cells sensitive to the *direction of virtual movement* of a train of click stimuli. Dichotically presented trains of clicks, having changing interaural time delays (simulating movement of a sound source), can evoke a response *only* if the direction of virtual movement is in a given direction, i.e. toward *or* away from the midline.

While, surprisingly, the role played by the colliculus in the localisation of sources of sound has not been established very clearly in experiments involving lesions therein, it may be responsible for the acoustic control of head and eye movements, mediated by the superior colliculus, in orientation towards sounds.

The *medial geniculate nucleus* appears to be organised in a similar manner to other thalamic nuclei, particularly the lateral geniculate and ventroposterior nuclei. The afferents, ascending from the inferior colliculus (Fig. 14.2), end in a small number of large terminals in aggregations on the dendrites of the geniculate cells of the ventral nucleus, the main 'through-pathway' to the cortex. The other divisions (medial and dorsal nuclei) receive input also from elsewhere than the inferior colliculus and are probably multimodal. They project predominantly to the secondary auditory cortex and other cortical areas.

Like the central nucleus of the inferior colliculus, the ventral nucleus is laminated, with a cochleotopic organisation. Low characteristic frequencies are represented in the outermost laminae and high in the innermost. Response areas are similar to those found at the cochlear nucleus level, with a high proportion of neurons being inhibited by tones. Intranuclear inhibition tends to emphasise the onset of stimuli and to curtail the sustained excitation seen at lower levels.

14.8. AUDITORY CORTEX

Organisation

The auditory cortex in primates is largely buried in the sylvian (lateral) fissure and located mainly on the inner surface of the superior temporal gyrus (Fig. 14.28*b*). In lower mammals such as the cat, it is conveniently spread out over a large part of the lateral surface of each hemisphere (Fig. 14.28*a*).

It is divided into a central core area, called primary cortex or AI, and various, surrounding, areas forming an auditory 'belt' (CM, L, RL in Fig. 14.28*b*; AII, Ep, SSF in Fig. 14.28*a*). AI receives the most direct, short-latency input from the medial geniculate nucleus. The surrounding belt areas, while also receiving input from the ventral division of the geniculate, receive projections from other thalamic nuclei. Consequently, some of the belt areas are polysensory areas.

The thalamic projections to most if not all these areas are each topographically organised. Each area contains a complete representation of the cochlea, although the scale and direction of the representation differs from area to area. In the primary area (AI) of the cat, the cochlear base is represented anteriorly and the apex posteriorly (Fig. 14.28*a*); in the primate, the representation is reversed (Fig. 14.28*b*). The basal region of the cochlea is relatively over-represented in AI, so that in cat and monkey over half of the

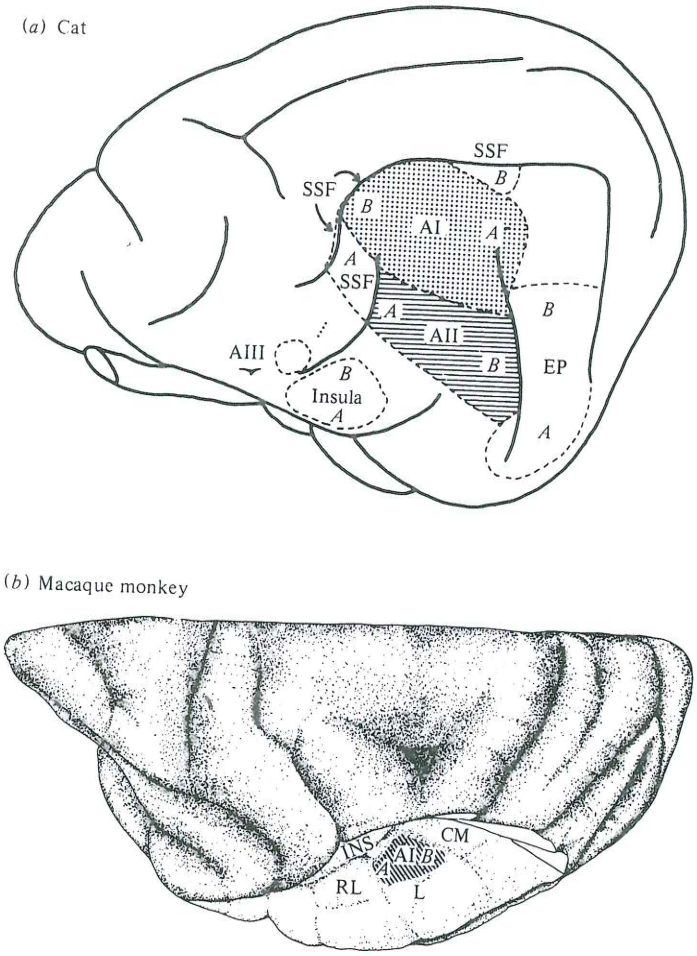
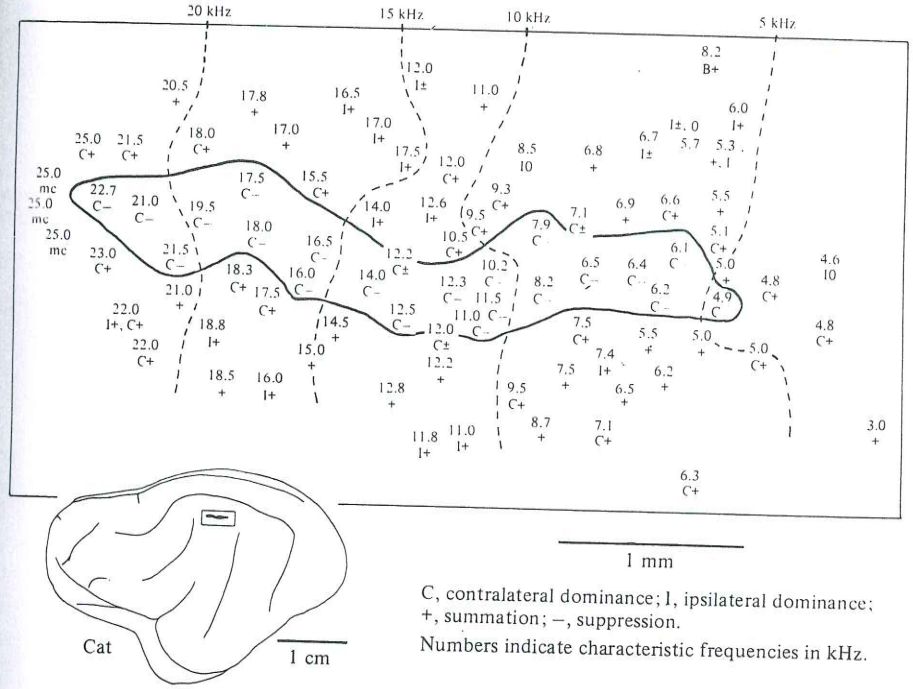


Fig. 14.28. Functional organisation of auditory cortex of monkey and cat. (*a*) Cat: view of left hemisphere from side. (*b*) Macaque monkey: view of left hemisphere from above. The parietal cortex has been removed to show auditory cortex spread out on the inner surface of the superior temporal plane. In both cases, the central primary cortex, AI, which receives the most direct input from the ventral nucleus of the medial geniculate nucleus, is surrounded by a 'belt' of cortex (AII, Ep, SSF, in (*a*); CM, L, RL, in (*b*). *A* and *B* indicates order of projection of fibres related to apical and basal ends of cochlea. (After Woolsey (1960). In *Neural Mechanisms of the Auditory and Vestibular Systems*, p. 165. Thomas; Brugge & Merzenich (1973). In *Basic Mechanisms in Hearing*, p. 745. Academic Press.)



C, contralateral dominance; I, ipsilateral dominance; +, summation; –, suppression.
Numbers indicate characteristic frequencies in kHz.

Fig. 14.29. Ear dominance 'slabs' in cat auditory cortex. Area (slab) enclosed in solid outline contains cells whose excitatory response is dominated by the contralateral ear, and suppressed by the ipsilateral ear. Areas above and below are summation cortex where the responses to the two ears are excitatory and summate, the contralateral (C) or ipsilateral (I) ears dominating the response. (From Imig & Adrian (1977). *Brain Research*, *138*, 241.)

area corresponds to cochlear regions of 10 kHz and above. In some species, substantial over-representation of frequencies of particular interest occurs, e.g. in the bat, for those concerned with echolocation. Intrinsic connections of the cortex, from one hemisphere to the other via the corpus callosum, are ordered so that homologous areas are interconnected. These connections end in cortical layer III, the afferent projections from the geniculate ending in layer IV. The primary area (AI) is reciprocally connected to some of the belt areas (AII, CM).

The primary area is also vertically organised into at least two overlapping systems of planes and slabs (Fig. 14.29), though this

organisation is not so clear-cut as in the visual and somatic cortical receiving areas. Lying at right angles to the cochleotopic axis of the primary area are planes, which extend the width of the area, each plane corresponding to a different locus of origin in the cochlea. Cortical cells in one of these planes tend to have similar characteristic frequencies. However, there is little sense in which the planes can be said to be frequency-specific: in unanaesthetised animals, a significant degree of variance of characteristic frequencies is encountered; many of the cells have broad response areas covering several octaves; and many of the cells are not responsive to pure tones at all. Orthogonal to these 'isofrequency' projection planes are slabs (e.g. solid outline in Fig. 14.29), which have a common binaural interaction pattern. There is anatomical evidence in the cat for the existence of two slabs in each hemisphere, where callosal fibres couple homologous areas of primary cortex. These are likely to correspond to the areas in Fig. 14.29 marked with a plus sign, where binaural *summation* occurs, i.e. the response to the binaural inputs are greater than those to the monaural inputs, although generally the contralateral or the ipsilateral ear dominates the response (C+ or I+). In the intervening areas where the callosal input is sparse, are situated the *suppression slabs*, one of which is shown in Fig. 14.29. Here, the contralateral input is dominant, the ipsilateral input exerting a suppressive effect on the cell's response to the contralateral input.

Selective response properties

A great variety of response types are encountered in the primary cortex, some of which are illustrated in Fig. 14.30 (A–F). In contrast to the situation at the auditory periphery, however, where all neurons respond to tones and wide-band signals (noise, clicks), not all cortical units respond to these unstructured stimuli. Half do not respond to wide-band noise in spite of having wide frequency response areas. On the other hand, the great majority can be activated by sound complexes structured in certain ways.

About 10% of cells will respond to tones only if their frequency is changing (Fig. 14.30*f*, *g*). Their response usually is to a preferred direction of frequency change: either up or down (Fig. 14.30*f*, *g*). In some cases it is to both directions of change. The rate of change of frequency and/or the repetition rate, if the modulation is periodic, are also important.

In addition to this selectivity for frequency change, direction and rate, certain cortical cells are selective for the temporal 'shape' of

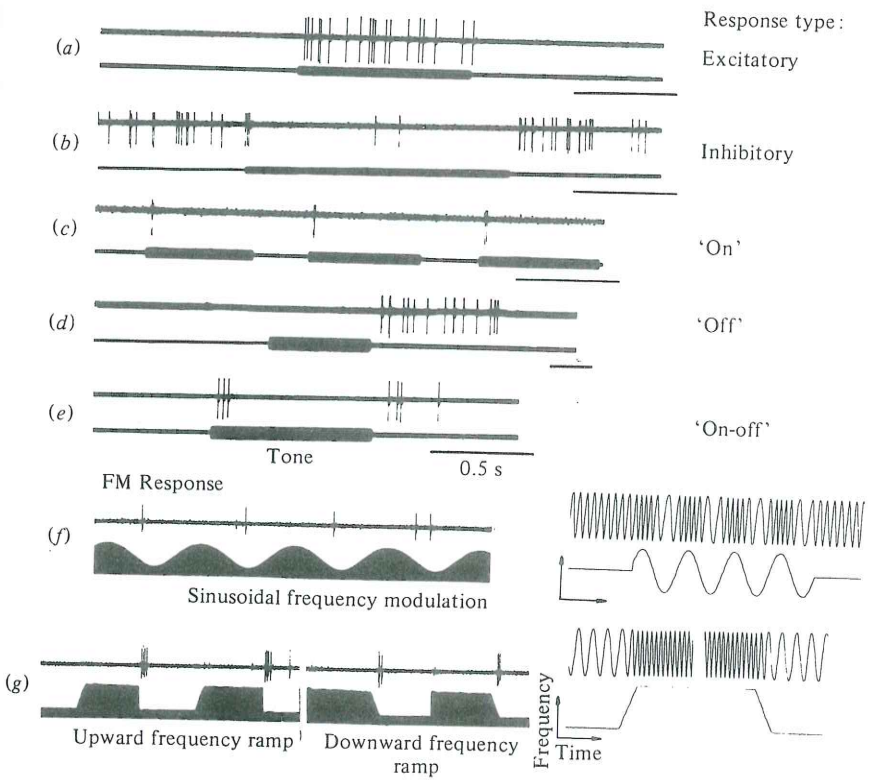


Fig. 14.30. Variety of response types of cells in primary auditory cortex. (*a*)–(*e*) Types of response to steady tones. (*f*, *g*) Cell responding to frequency-modulated, not steady tones. Note response selectivity to direction of frequency sweep: in the downward direction, not to upward sweeps. Black envelope indicates excursions of frequency as illustrated by the waveforms to the right. (Cat, after Evans & Whitfield (1964). *Journal of Physiology*, *171*, 476; Evans (1968) in *Hearing in Vertebrates*, p. 272. Churchill.)

the amplitude envelope of sounds. Others appear to be selective for noise or click stimuli and others for stimuli occurring within a narrow range of stimulus intensities.

In primates that have developed an extensive repertoire of socially meaningful vocalisations, such as squirrel monkeys, the cortical units are especially responsive to these calls. In fact, a small percentage of cells are selective for *individual* calls and remain unresponsive to other calls, tonal stimuli, noise, etc.

Some auditory cells, therefore, like their counterparts of other modalities (Chapter 1), can be said to be *abstractive* for certain features of complex stimuli which may have particular biological significance for the animal. Thus cortical cells are capable of providing 'answers' to the following questions: is the stimulus a noise, a click, a tone, or a species-specific call? Is the stimulus on? Has it just commenced? Has it just been terminated? What is its duration and repetition rate? Is the frequency changing? If so, in which direction; at what rate? How rapidly is the amplitude rising? Where is the stimulus located in space? Is it moving? And so on.

In certain bats, neurons selective for a particular feature are grouped in a restricted cortical area. These are units selective for changes in frequency that are characteristic of their calls used for echolocation. This is a highly specialised function, and may, like other such specialisations in the bat, not be generalisable to other mammals, where there is no discernible order to the location of these selectivities.

In man, there is psychophysical evidence for 'channels' specific for modulation at different frequencies. Exposure to repeated stimulation at a given modulation frequency or in one direction of frequency change, can raise the threshold of detection for those particular parameters. From similar experiments with speech sounds, there is evidence for feature selectivities in the human auditory system for certain of the acoustic cues important for the recognition of the parts of speech (§15.3).

Functional effects of cortical damage

It is perhaps not surprising, in view of the relative lack of specificity of cortical cells for frequency and intensity *per se*, that bilateral ablation of auditory cortex in cat and monkey leaves minimal deficits (upon relearning) in the behavioural audiogram, and in the discrimination and generalisation of frequency and intensity. The deficits are in the ability to localise or lateralise sounds, or distinguish changes in their duration, temporal pattern (order or sequence), or spectral composition. Animals relearn with difficulty to distinguish the direction of change of frequency. Similarly, in man, damage to the temporal lobe has been found to produce deficiencies in localisation in the contralateral sound field. However, in man, there is evidence of some hemispheric asymmetry: damage to the temporal lobe of the dominant hemisphere (the left in the case of a right-handed person) produces deficits predominantly in the recognition of speech; whereas damage to the non-dominant lobe affects recognition of the timbre of sounds, tonal sequence and pattern, i.e. some of the attributes of music.

14.9. DESCENDING, EFFERENT, PATHWAYS

Organisation

The ascending auditory pathway is accompanied, throughout its length, by a descending, efferent set of pathways (Fig. 14.1). One short pathway establishes reciprocal connections between the primary auditory cortex and the ventral nucleus of the medial geniculate nucleus. The second pathway runs from primary and surrounding cortical areas to the central and external nuclei of the inferior colliculus. From there it diverges to supply the dorsal cochlear nucleus, and the periolivary nuclei. The latter in turn give rise to the *olivocochlear bundle*. The ventral cochlear nucleus receives efferent connections from the lateral region of the superior olive.

Function

While the effects of electrically stimulating the efferent system are well known, paradoxically its action in the behaving animal is less clear. Stimulation of the efferent pathways can produce excitatory and inhibitory effects in cells of the cochlear nucleus, predominantly inhibition in the dorsal, and excitation in the ventral nuclei. However, stimulation of the olivocochlear bundle results only in *inhibition* of the evoked activity in cochlear fibres; this effect is mediated by acetylcholine release at the efferent endings in the cochlea. These inhibitory actions are frequency specific: they increase the fibre's threshold (by up to 25 dB in some fibres) predominantly at the characteristic frequency, thus reducing the sharpness of tuning. This efferent inhibition does not, however, affect significantly the spontaneous or the saturated discharge rates of fibres.

The fibres of the olivocochlear bundle itself are inactive in anaesthetised animals, and most exhibit no spontaneous activity in the absence of anaesthesia. However, in the latter case, tones evoke a slow regular discharge beginning after a relatively long latency of 5–40 ms.

In contrast to these clear effects of the olivocochlear bundle on cochlear activity, surgically interrupting the bundle appears to be without effect on auditory *behaviour*. Thus, no conclusive effects have

been found on absolute behavioural thresholds, thresholds in noise, frequency discrimination or selective attention. There is some evidence, however, that modification of efferent activity to the cochlear nucleus may affect an animal's ability to detect signals in noise. Inhibition of responses in the lateral lemniscus has been observed in behaving animals immediately before vocalisation, suggesting that one important – possibly the only – role of the efferent system (like that of the middle-ear muscle system) is to reduce activation of the auditory system in response to sounds generated by the animal itself.

14.10. EVOKED POTENTIALS FROM THE HUMAN AUDITORY SYSTEM

In clinical practice, useful information can be provided by recording electrical activity from the human auditory system without the invasive means required to investigate the response of single neurons in animals. The *electrocochleogram* has already been referred to (§14.3). For information on the electrical activity of more central structures sound-evoked potentials can be obtained (with an averaging computer) between an electrode simply placed on the scalp over the vertex of the head and one on the skin over the mastoid or on the ear lobe. At least 15 waves have been identified in response to click stimuli, and are conventionally labelled as in the diagrammatic representation (on a logarithmic time scale) of Fig. 14.31. Waves I to VI are the so-called *brain-stem evoked responses*, and some of them have been tentatively identified as indicating activity in a specific location, e.g. I: cochlear nerve; II: cochlear nucleus; III: superior olive; IV/V: inferior colliculus. The sites of origin of the so-called middle-latency waves, (N_0–N_b) and the long latency potentials (P_1–N_2) are very uncertain. The medial geniculate nucleus and the auditory cortex have been implicated in the generation of waves N_0 and N_a respectively. However, with the exception of the cochlear nerve, it is unlikely that the various waves can represent components *specific* to the nuclei, in view of the wide range of latencies encountered in the responses of individual neurons in each nucleus in animal experiments, from the cochlear nucleus onwards. It is possible that the waves represent the synchronous activity of those neurons with the shortest latency (predominantly found in the ventral acoustic pathway, p. 293) in the *nerve fibre tracts connecting the nuclei.*

The recording of these responses, particularly the early waves I to V, is proving to be a useful clinical tool. First, it provides information

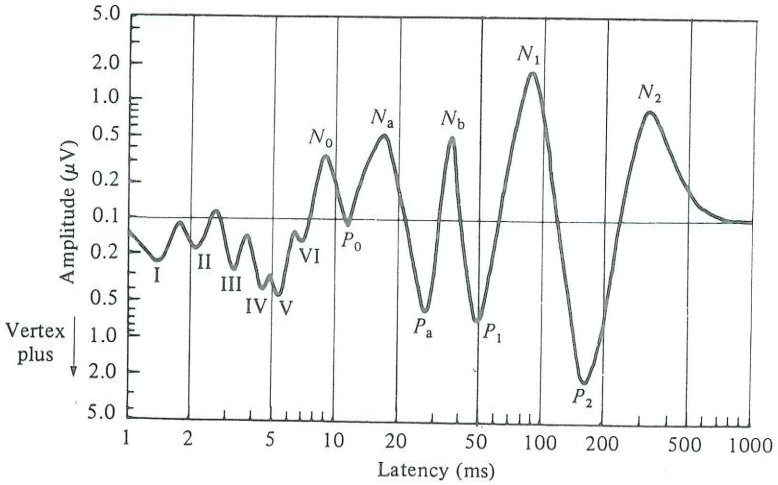


Fig. 14.31. Human auditory evoked potentials. Diagrammatic representation of the potentials obtained with the aid of an averaging computer), between an electrode on the vertex of the head and over the mastoid or on the ear lobe. Waves I to VI are called the short-latency or brain stem evoked responses; waves N_0 to N_b the middle-latency responses, and waves P_1 to N_2 the late-latency responses. (From Picton *et al.* (1974). *Electroencephalography and clinical Neurophysiology*, *36*, 179.)

of value for localising lesions within the central auditory pathways. Second, it allows objective measurements of hearing threshold to be made without invasive procedures in non-communicating children and adults. The long-latency waves are markedly affected by the state of arousal or sedation, and are therefore less useful clinically. However, the longest latency waves (e.g. N_2 and later) are systematically affected by cognitive processes and may turn out to be of value in the investigation of language disorders.

14.11. SUGGESTIONS FOR FURTHER READING

General introductions and reviews

Yost, W. A. & Nielsen, D. W. (1977) *Fundamentals of Hearing*. New York: Holt, Rinehart & Winston. (Well illustrated, clear introduction to the peripheral auditory system.)

Durrant, J. D. & Lovrinic, J. H. (1977) *Bases of Hearing Science*. Baltimore: Williams & Wilkins Co.

Special references

Outer, middle and inner ear. Dallos, P. (1973) *The Auditory Periphery.* New York: Academic Press. Relevant chapters in *Handbook of Sensory Physiology*, vol. 5, pt 1 (1974), ed. W. D. Keidel & W. D. Neff. Berlin: Springer-Verlag.

Sharp tuning of basilar membrane. Sellick, P. M., Patuzzi, R. & Johnstone, B. M. (1982) *Journal of the Acoustical Society of America*, **72**, 131–41; Khanna, S. M. & Leonard, D. G. B. (1982) *Science*, **215**, 305–6.

Cochlear hair cells. Russell, I. J. & Sellick, P. M. (1978) *Journal of Physiology*, **284**, 261.

Electrical resonance of hair cells. Crawford, A. C. & Fettiplace, R. (1981) *Journal of Physiology*, **312**, 377.

Cochlear nerve and nucleus. Kiang, N. Y-S., Watanabe, T., Thomas, E. C. & Clarke, L. F. (1965) *Discharge patterns of single fibers in the cat's auditory nerve.* Cambridge, Mass: MIT Press. Evans, E. F. (1975) Cochlear nerve and cochlear nucleus. In *Handbook of Sensory Physiology*, vol. 5, pt 2, ed. W. D. Keidel & W. D. Neff. Berlin: Springer-Verlag.

Superior olive, medial geniculate nucleus and cortex. Relevant chapters in *Handbook of Sensory Physiology*, vol. 5, pt 2, ed. W. D. Keidel & W. D. Neff (1975) Berlin: Springer-Verlag.

Feature- and call-specific neurons in auditory pathway. Evans, E. F. (1974) In *The Neurosciences: Third Study Program*, ed. F. C. Schmitt & F. G. Worden. Cambridge, Mass: MIT Press.

Binaural columns in primary auditory cortex. Imig, T. J. & Adrian, H. O. (1977) *Brain Research*, **138**, 241–57.

CHAPTER FIFTEEN

Functions of the auditory system

E. F. EVANS

15.1. PROCESSING OF FREQUENCY

Frequency analysis

If the individual components of a complex sound are sufficiently far apart in frequency, they can be heard as separate, as in the notes of a musical chord. This is *Ohm's acoustic law*. Closer together they are heard as one sound. The ability of the ear to resolve sounds into their constituent (*simultaneously* present) frequency components, is termed *frequency analysis or frequency selectivity*. It is essential if the ear is to be able to follow simultaneous changes in the frequency and intensity of the individual components of a complex sound, as, for example, in speech sounds (Fig. 13.1*h–l*).

In many ways, the ear acts as if it contains a bank of narrowly tuned filters. Psychophysical estimates of the shapes of these filters (as in Fig. 15.1*a*) resemble those of the filters represented by the frequency threshold curves of cochlear nerve fibres (as in Fig. 14.16), and many of the frequency-selective properties of the ear can be accounted for (to a first approximation) on this basis. To understand some of the ear's frequency selective properties, it is helpful to consider these neural and psychophysical frequency threshold 'tuning' curves as representing linear filter functions, as in Fig. 15.1*b*. Here the curve for a single cochlear fibre is drawn, inverted, on *linear* power and frequency scales. It has a central '*pass band*' about the characteristic frequency (6.4 kHz) and (on this scale), nearly symmetrical '*skirts*' or filter cut-offs (for the first 10 dB or so). In a linear filter, the width of the pass band can be usefully expressed as the *effective bandwidth* (Fig. 15.1*b*). This is the bandwidth of the rectangular filter having the same area; it is approximately the half-power (3 dB down) bandwidth. The concept of effective bandwidth is important because with multicomponent or broad-band signals, it is the frequencies falling within the effective bandwidth that dominate the filter's response, compared with those falling outside.