Chapter 2
Auditory Periphery: From Pinna to Auditory Nerve

Ray Meddis and Enrique A. Lopez-Poveda

Abbreviations and Acronyms

AC	Alternating current
AN	Auditory nerve
BF	Best frequency
BM	Basilar membrane
BW	Bandwidth
CF	Characteristic frequency
dB	Decibel
DC	Direct current
DP	Distortion product
DRNL	Dual-resonance nonlinear
f_C	Center frequency
FFT	Fast Fourier transform
FIR	Finite impulse response
HRIR	Head-related impulse response
HRTF	Head-related transfer function
HSR	High-spontaneous rate
IHC	Inner hair cell
IIR	Infinite impulse response
kHz	KiloHertz
LSR	Low-spontaneous rate
MBPNL	Multiple bandpass nonlinear
ms	Milliseconds
OHC	Outer hair cell
SPL	Sound pressure level

R. Meddis (✉)
Hearing Research Laboratory, Department of Psychology, University of Essex, Colchester CO4 3SQ, UK
e-mail: rmeddis@essex.ac.uk

R. Meddis et al. (eds.), *Computational Models of the Auditory System*,
Springer Handbook of Auditory Research 35, DOI 10.1007/978-1-4419-5934-8_2,

2.1 Introduction

The auditory periphery begins at the point where the pressure wave meets the ear and it ends at the auditory nerve (AN). The physical distance is short but the sound is transformed almost beyond recognition before it reaches the end of its journey. The process presents a formidable challenge to modelers, but considerable progress has been made over recent decades.

The sequence starts as a pressure wave in the auditory meatus, where it causes vibration of the eardrum. These vibrations are transmitted to the stapes in the middle ear and then passed on to the cochlear fluid. Inside the cochlea, the basilar membrane (BM) responds with tuned vibrations that are further modified by neighboring outer hair cells (OHCs). This motion is detected by inner hair cells (IHCs) that transduce it into fluctuations of an electrical receptor potential that control indirectly the release of transmitter substance into the AN synaptic cleft. Finally, action potentials are generated in the tens of thousands of auditory nerve fibers that carry the auditory message to the brain stem. Each of these successive transformations contributes to the quality of hearing, and none can be ignored in a computer model of auditory peripheral processing.

This combined activity of processing stages is much too complex to be understood in an intuitive way, and computer models have been developed to help us visualize the succession of changes between the eardrum and the AN. The earliest models used analogies with electrical tuned systems such as radio or radar, and these continue to influence our thinking. However, the most recent trend is to simulate as closely as possible the individual physiological processes that occur in the cochlea. Model makers are guided by the extensive observations of anatomists and physiologists who have mapped the cochlea and measured the changes that occur in response to sound. Their measurements are made at a number of places along the route and include the vibration patterns of the eardrum, stapes, and BM; the electrical potentials of the OHCs and IHCs; and, finally, the action potentials in the AN fibers. These places mark “way points” for modelers who try to reproduce the physiological measurements at each point. Successful simulation of the physiological observations at each point is the main method for verifying their models. As a consequence, most models consist of a cascade of “stages” with the physiological measurement points marking the boundary between one stage and another. The freedom to model one stage at a time has greatly simplified what would otherwise be an impossibly complex problem.

Figure 2.1 illustrates a cascade model based on the work conducted by the authors. The signal is passed from one stage to another, and each stage produces a unique transformation to simulate the corresponding physiological processes. Two models are shown. On the left is a model of the response at a single point along the BM showing how the stapes displacement is transformed first into BM displacement, then into the IHC receptor potential, and then into a probability that a vesicle of transmitter will be released onto the IHC/AN synaptic cleft (if one is available). The bottom panel shows the spiking activity of a number of auditory

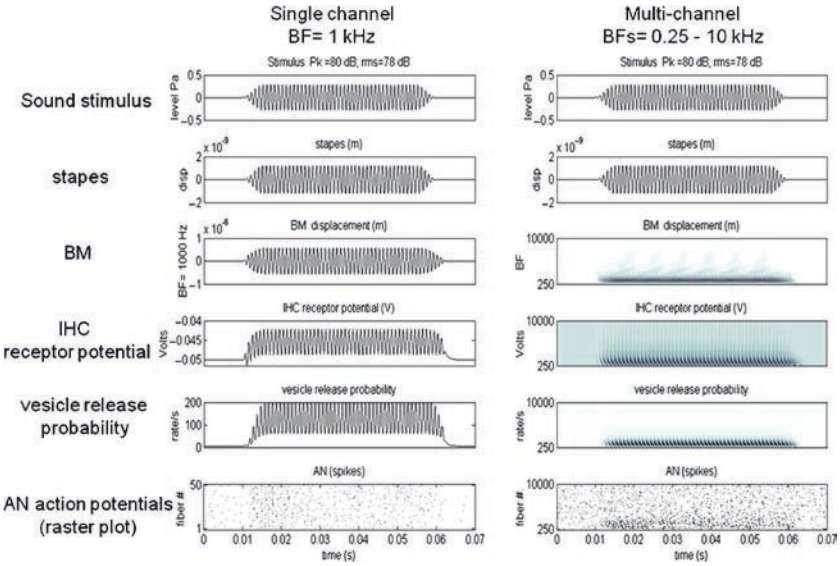

Fig. 2.1 The response of a multistage computer model of the auditory periphery is illustrated using a 1-kHz pure tone presented for 50 ms at 80 dB SPL. Each *panel* represents the output of the model at a different stage between the stapes and the auditory nerve. The *left-hand panels* show a single channel model (BF = 1 kHz) representing the response at a single point along the basilar membrane. Each plot shows the response in terms of physical units: stapes (displacement in meters), the BM (displacement in meters), the IHC receptor potential (volts), and vesicle release (probability). The *right-hand panels* show surface plots representing the response of a 40-channel model with BFs ranging between 250 Hz and 10 kHz. Channels are arranged across the *y*-axis (high BFs at the *top*) with time along the *x*-axis. *Darker shading* indicates more activity. Note that high-BF channels are only weakly affected by the 1-kHz pure tone and most activity is concentrated in the low-BF channels. The *bottom panel* of both models is the final output of the model. It shows the spiking activity of a number of AN fibers represented as a raster plot where each *row of dots* is the activity of a single fiber and each *dot* is a spike. The *x*-axis is time. In the single-channel model (*left*), all fibers have the same BF (1 kHz). In the multichannel model (*right*), the fibers are arranged with high-BF fibers at the *top*. Note that all fibers show spontaneous activity and the response to the tone is indicated only by an increase in the firing rate, particularly at the beginning of the tone. In the multichannel model, the *dots* can be seen to be more closely packed in the low-BF fibers during the tone presentation

nerve fibers presented as a raster plot where each dot represents a spike in a nerve fiber. On the right, a more complex model is shown. This represents the activity at 40 different sites along the cochlear partition each with a different best-frequency (BF). Basal sites (high BFs) are shown at the top of each panel and apical sites (low BF) at the bottom with time along the x-axis. Darker shades indicate more intense activity.

The input to the model is a 1-kHz ramped tone presented for 50 ms at a level of 80 dB SPL. The multichannel model shows frequency selectivity in that only some channels are strongly affected by the stimulus. It is also important to note that the AN fibers are all spontaneously active, and this can be seen most clearly before the tone begins to play. The single-channel model (left) shows most frequent firing soon after the onset of the tone, and this is indicated by more closely packed dots in the raster plot. When the tone is switched off, the spontaneous firing is less than before the tone, as a consequence of the depletion of IHC presynaptic transmitter substance that has occurred during the presentation of the tone. The multichannel model (right) shows a substantial increase of AN fiber firing only in the apical channels (low-BFs at the bottom of the plot). Only a small number of fibers are shown in the figure to illustrate the basic principles. A full model will represent the activity of thousands of fibers.

Models serve many different purposes, and it is important to match the level of detail to the purpose in hand. For example, psychophysical models such as the loudness model of Moore et al. (1997) are based only loosely on physiology including a preemphasis stage (outer–middle ear), as well as frequency tuning and compression (BM). When compared with the model in Fig. 2.1, it is lacking in physiological detail. Nevertheless, it serves an important purpose in making useful predictions of how loud sounds will appear to the listener. When fitting hearing aids, for example, this is very useful and the model is fit for its purpose. By contrast, the more detailed simulations of the auditory periphery (discussed in this chapter) cannot at present make loudness predictions.

A more detailed model such as that offered by Derleth et al. (2001) includes peripheral filtering and a simulation of physiological adaptation without going so far as to model the individual anatomical components. This has proved useful in simulating human sensitivity to amplitude modulation. It may yet prove to be the right level of detail for low-power hardware implementations such as hearing aids because the necessary computing power is not available in a hearing aid to model all the details of a full physiological model. Different degrees of detail are required for different purposes. Nevertheless, in this chapter, emphasis is placed on computer models that simulate the anatomy and physiology as closely as possible because these are the only models that can be verified via actual physiological measurements.

Auditory models can be used in many different ways. From a purely scientific point of view, the model represents a theory of how the auditory periphery works. It becomes a focus of arguments among researchers with competing views of the underlying "truth." In this respect, computer models have the advantage of being quantitatively specified because their equations make quantitative predictions that can be checked against the physiological data. However, models also have the potential for practical applications. Computer scientists can use a peripheral model as an input to an automatic speech recognition device in the hope that it will be better than traditional signal-processing methods. Such attempts have had mixed success so far but some studies have found this input to be more robust (Kleinschmidt et al. 1999). Another application involves their use in the design of algorithms for generating the signals used in cochlear implants or hearing aids (e.g., Chapter 9; Chapter 7). Indeed, any problem involving the analysis of acoustic signals might benefit from the use of auditory models, but many of these applications lie in the future.

Before examining the individual stages of peripheral auditory models, some preliminary remarks are necessary concerning the nature of compression or "nonlinearity" because it plays an important role in many of these stages. In a linear system, an increase in the input signal results in a similar-size increase at the output; in other words, the level of the output can be predicted as the level of the input multiplied by a constant. It is natural to think of the auditory system in these terms. After all, a sound is perceived as louder when it becomes more intense. However, most auditory processing stages respond in a nonlinear way. The vibrations of the BM, the receptor potential in the IHC, the release of transmitter at the IHC synapse, and the auditory nerve firing rate are all nonlinear functions of their inputs. The final output of the system is the result of a cascade of nonlinearities. Such systems are very difficult to intuit or to analyze using mathematics. This is why computer models are needed. This is the only method to specify objectively and test how the system works.

The auditory consequences of this compression are important. They determine the logarithmic relationship between the intensity of a pure tone and its perceived intensity. It is for this reason that it is important to describe intensity using decibels rather than Pascals when discussing human hearing. Further, when two tones are presented at the same time they can give rise to the perception of mysterious additional tones called "combination tones" (Goldstein 1966; Plomp 1976). The rate of firing of an auditory nerve in response to a tone can sometimes be reduced by the addition of a second tone, known as two-tone suppression (Sachs and Kiang 1968). The width of an AN "tuning curve" is often narrow when evaluated near threshold but becomes wider when tested at high signal levels. These effects are all the emergent properties of a complex nonlinear system. Only computer models can simulate the consequences of nonlinearity, especially when complex broadband sounds such as speech and music are being studied.

The system is also nonlinear in time. The same sound produces a different response at different times. A brief tone that is audible when presented in silence may not be audible when it is presented after another, more intense tone, even though a silent gap may separate the two. The reduction in sensitivity along with the process of gradual recovery is known as the phenomenon of "adaptation" and it is important to an understanding of hearing in general. Once again, this nonlinearity can be studied effectively only by using computer simulation.

This chapter proceeds, like a peripheral model, by examining each individual processing stage separately and ending with the observation that the cascade of stages is complicated by the presence of feedback loops in the form of the efferent system that has only recently began to be studied. Finally, some examples of the output of a computer model of the auditory periphery are evaluated.

2.2 Outer Ear

The first stage of a model of the auditory periphery is the response of the middle ear, but it must be remembered that sounds are modified by the head and body of the listeners before they enter the ear canal. In a free-field situation, the spectrum

of a sound is first altered by the filtering action of the body (Shaw 1966; Lopez-Poveda 1996). The acoustic transfer function of the body in the frequency domain is commonly referred to as the head-related transfer function (HRTF) to stress that the principal filtering contributions come from the head and the external ear (Shaw 1975; Algazi et al. 2001). In the time domain, the transfer function is referred to as the head-related impulse response (HRIR). The HRIR is usually measured as the click response recorded by either a miniature microphone placed in the vicinity of the eardrum (Wightman and Kistler 1989) or by the microphone of an acoustic manikin (Burkhard and Sachs 1975). The filtering operation of the body is linear; thus a Fourier transform serves to obtain the HRTF from its corresponding HRIR.

The spectral content of an HRTF reflects diffraction, reflection, scattering, resonance, and interference phenomena that affect the incoming sound before it reaches the eardrum (Shaw 1966; Lopez-Poveda and Meddis 1996). These phenomena depend strongly on the location of the sound source relative to the ear's entrance, as well as on the size and shape of the listener's torso, head, pinnae, and ear canal. As a result, HRTFs, particularly their spectral characteristics above 4 kHz, are different for different sound source locations and for different individuals (Carlile and Pralong 1994). Further, for any given source location and individual, the HRTFs for the left and the right ear are generally different as a result of the two ears being slightly dissimilar in shape (Searle et al. 1975). The location-dependent spectral content of HRTFs is a useful cue for sound localization, and for this reason HRTFs have been widely studied (Carlile et al. 2005).

2.2.1 Approaches to Modeling the Head-Related Transfer Function

All of the aforementioned considerations should give an idea of the enormous complexity involved in producing a computational model of HRTFs. Nevertheless, the problem has been attempted from several angles. There exists one class of models that try to reproduce the main features of the HRTFs by mathematically formulating the physical interaction of the sound waves with the individual anatomical elements of the body. For example, Lopez-Poveda and Meddis (1996) reproduced the elevation-dependent spectral notches of the HRTFs considering that the sound is diffracted at the concha aperture and then reflected on the concha back wall before reaching the ear canal entrance. The total pressure at the ear canal entrance would be the sum of the direct sound plus the diffracted/reflected sound. Similar physical models have been developed by Duda and Martens (1998) to model the response of a spherical head, by Algazi et al. (2001) to model the combined contributions of a spherical head and a spherical torso, and by Walsh et al. (2004) to model the combined contribution of the head and the external ear.

One of the main advantages of physical models is that they help elucidate the contributions of the individual anatomical elements to the HRTFs. Another advantage is that they allow approximate HRTFs to be computed for (theoretically) arbitrary

body geometries, given the coordinates of the sound source(s). In practice, however, they are usually evaluated for simplified geometrical shapes (an exception is the model of Walsh et al. 2004) and are computationally very expensive. Another disadvantage is that, almost always, these models are developed in the frequency domain, although the HRIR can be obtained from the model HRTF by means of an inverse Fourier transform (Algazi et al. 2001). For these reasons, physical models of HRTFs are of limited practical use as part of composite models of spectral processing by the peripheral auditory system.

An alternative method is to reproduce specific HRTFs by means of finite- (FIR) or infinite-impulse response (IIR) digital filters. An immediately obvious way to approach it is to treat the sample values of the experimental digital HRIRs as the coefficients of an FIR filter (Kulkarni and Colburn 2004). Alternatively, such coefficients may be obtained by an inverse Fourier transform of the amplitude HRTF (e.g., Lopez-Poveda and Meddis 2001), although this method does not preserve the phase spectra of HRIRs that may be perceptually important (Kulkarni et al. 1999).

A more challenging problem, however, is to develop computationally efficient digital filter implementations of HRIRs, that is, digital filters of the lowest possible order that preserve the main amplitude and phase characteristics of the HRTFs. This is important to obtain HRIRs that can be computed in real time. The problem is twofold. First, it is necessary to identify the main spectral characteristics of HRTFs that are common to all individuals and provide important sound localization information (Kistler and Wightman 1992). Second, it is necessary to reproduce those features using low-order IIR filters, as they are more efficient than FIR filters. Kulkarni and Colburn (2004) have recently reported a reasonable solution to the problem by demonstrating that stimuli rendered through a 6-pole, 6-zero IIR-filter model of the HRTF had inaudible differences from stimuli rendered through the actual HRTF.

The main advantages of these digital-filter-type models is that they can process time-varying signals in real or quasi-real time. Their disadvantages are that they shed no light on the physical origin or the anatomical elements responsible for the characteristic spectral features of the HRTFs. Further, they require that the HRTFs of interest be measured beforehand (several publicly available databases already exist). Nevertheless, this type of model is more frequently adopted in composite models of signal processing by the peripheral auditory system.

2.3 Middle Ear

The middle ear transmits the acoustic energy from the tympanic membrane to the cochlea through a chain of three ossicles: the malleus, in contact with the eardrum, the incus, and the stapes, which contacts the cochlea at the oval window. The middle ear serves to adapt the low acoustic impedance of air to that of the cochlear perilymphatic fluid, which is approximately 4,000 times higher (von Helmholtz 1877; Rosowski 1996). For frequencies below approximately 2 kHz, this impedance transformation is accomplished mainly by the piston-like functioning of the middle ear (Voss et al. 2000)

that results from the surface area of the eardrum being much larger than that of the stapes footplate. The lever ratio of the ossicles also contributes to the impedance transformation for frequencies above approximately 1 kHz (Goode et al. 1994).

In signal processing terms, the middle ear may be considered as a linear system whose input is a time-varying pressure signal near the tympanic membrane, and whose corresponding output is a time-varying pressure signal in the scala vestibuli of the cochlea, next to the stapes footplate. Therefore, its transfer function is expressed as the ratio (in decibels) of the output to the input pressures as a function of frequency (Nedzelnitsky 1980; Aibara et al. 2001). The intracochlear pressure relates directly to the force exerted by the stapes footplate, which in turn relates to the displacement of the stapes with respect to its resting position. For pure tone signals, stapes velocity (v) and stapes displacement (d) are related as follows: $v = 2\pi fd$, where f is the stimulus frequency in Hertz. For this reason, it is also common to express the frequency transfer function of the middle ear as stapes displacement or stapes velocity vs. frequency for a given sound level (Goode et al. 1994).

The middle ear is said to act as a linear system over a wide range of sound levels (<130 dB SPL) for two reasons. First, the intracochlear peak pressure at the oval window (Nedzelnitsky 1980), the stapes peak displacement (Guinan and Peake 1966), or the stapes peak velocity (Voss et al. 2000) is proportional to the peak pressure at the eardrum. The second reason is that sinusoidal pressure variations at the tympanic membrane produce purely sinusoidal pressure variations at the oval window (Nedzelnitsky 1980). In other words, the middle ear does not introduce distortion for sound levels below approximately 130 dB SPL.

The middle ear shapes the sound spectrum because it acts like a filter. However, a debate has been recently opened on the type of filter. Recent reports (Ruggero and Temchin 2002, 2003) suggest that the middle ear is a wide-band pressure transformer with a flat velocity-response function rather than a bandpass pressure transformer tuned to a frequency between 700 and 1,200 Hz, as previously thought (Rosowski 1996). The debate is still open.

2.3.1 Approaches to Modeling the Middle Ear Transfer Function

The function of the middle ear has been classically modeled by means of analog electrical circuits (Møller 1961; Zwislocki 1962; Kringlebotn 1988; Goode et al. 1994; Pascal et al. 1998; Voss et al. 2000; reviewed by Rosowski 1996). These models regard the middle ear as a transmission line with lumped mechanical elements and, as such, its functioning is described in electrical terms thanks to the analogy between electrical and acoustic elements (this analogy is detailed in Table 2.2 of Rosowski 1996). These models commonly describe the middle ear as a linear filter, although the model of Pascal et al. (1998) includes the nonlinear effects induced by the middle-ear reflex that occur at very high levels (>100 dB SPL). Electrical analogues have also been developed to model the response of pathological (otosclerotic) middle ear function (Zwislocki 1962).

The function of the middle ear has also been modeled by means of biomechanical, finite element methods (e.g., Gan et al. 2002; Koike et al. 2002; reviewed by Sun et al. 2002). This approach requires reconstructing the middle ear geometry, generally from serial sections of frozen temporal bones. The reconstruction is then used to develop a finite-element mesh description of the middle ear mechanics. So far, the efforts have focused on obtaining realistic descriptions of healthy systems that include the effects of the attached ligaments and tendons. However, as noted by Gan et al. (2002), finite element models will be particularly useful to investigate the effects of some pathologies (e.g., tympanic perforations or otosclerosis) on middle ear transmission, as well as to design and develop better middle ear prostheses (Dornhoffer 1998). These models also allow detailed research on the different modes of vibration of the tympanic membrane (e.g., Koike et al. 2002), which influence middle ear transmission for frequencies above approximately 1 kHz (Rosowski 1996). The main drawback of finite element models is that they are computationally very expensive.

A third approach is that adopted by most signal processing models of the auditory periphery. It consists of simulating the middle ear function by a linear digital filter with an appropriate frequency response. As a first approximation, some studies (e.g., Lopez-Poveda 1996; Robert and Eriksson 1999; Tan and Carney 2003) have used a single IIR bandpass filter while others (Holmes et al. 2004; Sumner et al. 2002, 2003a, b) use a filter cascade in an attempt to achieve more realistic frequency response characteristics. In any case, the output signal must be multiplied by an appropriate scalar to achieve a realistic gain.

Some authors have suggested that the frequency response of the middle ear determines important characteristics of the basilar response, such as the asymmetry of the iso-intensity response curves (Cheatham and Dallos 2001; see later) or the characteristic frequency modulation of basilar membrane impulse responses, that is, the so-called "glide" (e.g., Tan and Carney 2003; Lopez-Najera et al. 2005). This constitutes a reasonable argument in favor of using more realistic middle ear filter functions as part of composite models of the auditory periphery. To produce such a filters, some authors (e.g., Lopez-Poveda and Meddis 2001) employ FIR digital filters whose coefficients are obtained as the inverse fast Fourier transform (FFT) of an experimental stapes frequency response curve, whereas others (e.g., Lopez-Najera et al. 2007) prefer to convolve the tympanic pressure waveform directly with an experimental stapes impulse response. The latter approach guarantees realistic amplitude and phase responses for the middle ear function in the model.

2.4 Basilar Membrane

The motion of the stapes footplate in response to sound creates a pressure gradient across the cochlear partition that sets the organ of Corti to move in its transverse direction. The characteristics of this motion are commonly described in terms of BM velocity or displacement with respect to its resting position.

The BM responds tonotopically to sound. The response of each BM site is strongest for a particular frequency (termed the best frequency or BF) and decreases gradually with moving the stimulus frequency away from it. For this reason, each BM site is conveniently described to function as a frequency filter and the whole BM as a bank of overlapping filters. Each BM site is identified by its characteristic frequency (CF), which is defined as the BF for sounds near threshold.

BM filters are nonlinear and asymmetric. They are asymmetric in that the magnitude of the BM response decreases faster for frequencies above the BF than for frequencies below it as the stimulus frequency moves away from the BF (e.g., Robles and Ruggero 2001). The asymmetry manifests also in that the impulse (or click) response of a given BM site is modulated in frequency. This phenomenon is sometimes referred to as the chirp or glide of BM impulse responses. For basal sites, the instantaneous frequency of the impulse response typically increases with increasing time (Recio et al. 1998). The direction of the chirp for apical sites is still controversial (e.g., Lopez-Poveda et al. 2007), but AN studies suggest it could happen in the direction opposite to that of basal sites (Carney et al. 1999).

Several phenomena demonstrate the nonlinear nature of BM responses (Robles and Ruggero 2001). First, BM responses show more gain at low than at high sound levels. As a result, the magnitude of the BM response grows compressively with increasing sound level (slope of ~0.2 dB/dB). BM responses are linear (slope of 1 dB/dB) for frequencies an octave or so below the CF. This frequency response pattern, however, is true for basal sites only. For apical sites (CFs below ~1 kHz), compressive responses appear to extend to a wider range of stimulus frequencies relative to the CF (Rhode and Cooper 1996; Lopez-Poveda et al. 2003).

BM responses are nonlinear also because the BF and the bandwidth of a given cochlear site change depending on the stimulus level. The BF of basal sites decreases with increasing sound level. There is still controversy on the direction of change of the BF of apical cochlear sites. AN studies suggest that it increases with increasing level (Carney et al. 1999), but psychophysical studies suggest a downward shift (Lopez-Poveda et al. 2007). The bandwidth is thought to increase always with increasing level.

Suppression and distortion are two other important phenomena pertaining to BM nonlinearity (reviewed in Lopez-Poveda 2005). Suppression occurs when the magnitude of BM response to a given sound, called the suppressee, decreases in the presence of a second sound, called the suppressor. It happens only for certain combinations of the frequency and level of the suppressor and the suppressee (Cooper 1996, 2004). Suppression leads to decreases in both the degree (i.e., the slope) and dynamic range of compression that can be observed in the BM response. The time course of the two-tone suppression appears to be instantaneous (Cooper 1996).

Distortion can occur for any stimulus but is more clearly seen when the BM is stimulated with pairs of tones of different frequencies (f_1 and f_2, $f_2 > f_1$) referred to as primaries. In response to tone pairs, the BM excitation waveform contains distortion products (DPs) with frequencies $f_2 - f_1$, $(n+1)f_1 - nf_2$ and $(n+1)f_2 - nf_1$ ($n = 1, 2, 3, \ldots$) (Robles et al. 1991). These DPs are generated at cochlear sites with CFs equal to the primaries but can travel along the cochlea and excite remote BM regions with CFs equal to the DP frequencies (Robles et al. 1997). DPs can be heard as combination

tones (Goldstein 1966) and are thought to be the source of distortion-product otoacoustic emissions.

The characteristics of BM responses are not steady. Instead, they change depending on the activation of the efferent cochlear system, which depends itself on the characteristics of the sound being presented in the ipsilateral and contralateral ears. Activation of the efferent system reduces the cochlear gain (Russell and Murugasu 1997).

BM responses depend critically on the physiological state of the cochlea. Some diseases or treatments with ototoxic drugs (furosemide, quinine, aminoglycosides) damage cochlear outer hair cells, reducing the gain and the tuning of BM responses. Responses are fully linear postmortem or in cochleae with total OHC damage (reviewed in Ruggero et al. 1990; Robles and Ruggero 2001). Consequently, BM responses are sometimes described as the sum of an active (nonlinear) component, present only in cochleae with remaining OHCs, and a passive (linear) component, which remains post-mortem.

The BM response characteristics described in the preceding text determine important physiological properties of the AN response as well as perceptual properties in normal-hearing listeners and in those with cochlear hearing loss (Moore 2007). To a first approximation they determine, for instance, the frequency tuning of AN fibers near threshold (Narayan et al. 1998), the dynamic range of hearing (reviewed in Bacon 2004), our ability (to a limited extent) to resolve the frequency components of complex sounds (reviewed in Moore 2007), and even our perception of combination tones not present in the acoustic stimulus (Goldstein 1966). In addition, suppression is thought to facilitate the perception of speech immersed in certain kinds of noise (Deng and Geisler 1987; Chapter 9). Therefore, it is fundamental that composite AN models and models of auditory perception include a good BM nonlinear model.

2.4.1 Phenomenological BM Models

BM models aim at simulating BM excitation (velocity or displacement) in response to stapes motion. Many attempts have been made to achieve this with models of different nature. We review only a small a selection of phenomenological, signal-processing models. These types of models attempt to account for BM responses using signal-processing elements (e.g., digital filters). The advantage of this approach is that the resulting models can be implemented and evaluated easily for digital, time-varying signals. Models of a different kind are reviewed elsewhere: a succinct review of transmission line models is provided by Duifhuis (2004) and van Schaik (Chapter 10); mechanical cochlear models are reviewed by de Boer (1996). A broader selection of phenomenological models is reviewed in Lopez-Poveda (2005).

2.4.1.1 The MBPNL Model

The Multiple BandPass NonLinear (MBPNL) model of Goldstein (1988, 1990, 1993, 1995) was developed in an attempt to provide a unified account of complex BM nonlinear phenomena such as compression, suppression, distortion, and simple-tone

interference (the latter phenomenon is described later). It simulates the filtering function of a given cochlear partition (a given CF) by cascading a narrowly tuned bandpass filter followed by a compressive memoryless nonlinear gain, followed by another more broadly tuned bandpass filter (Fig. 2.2a). This structure is similar to

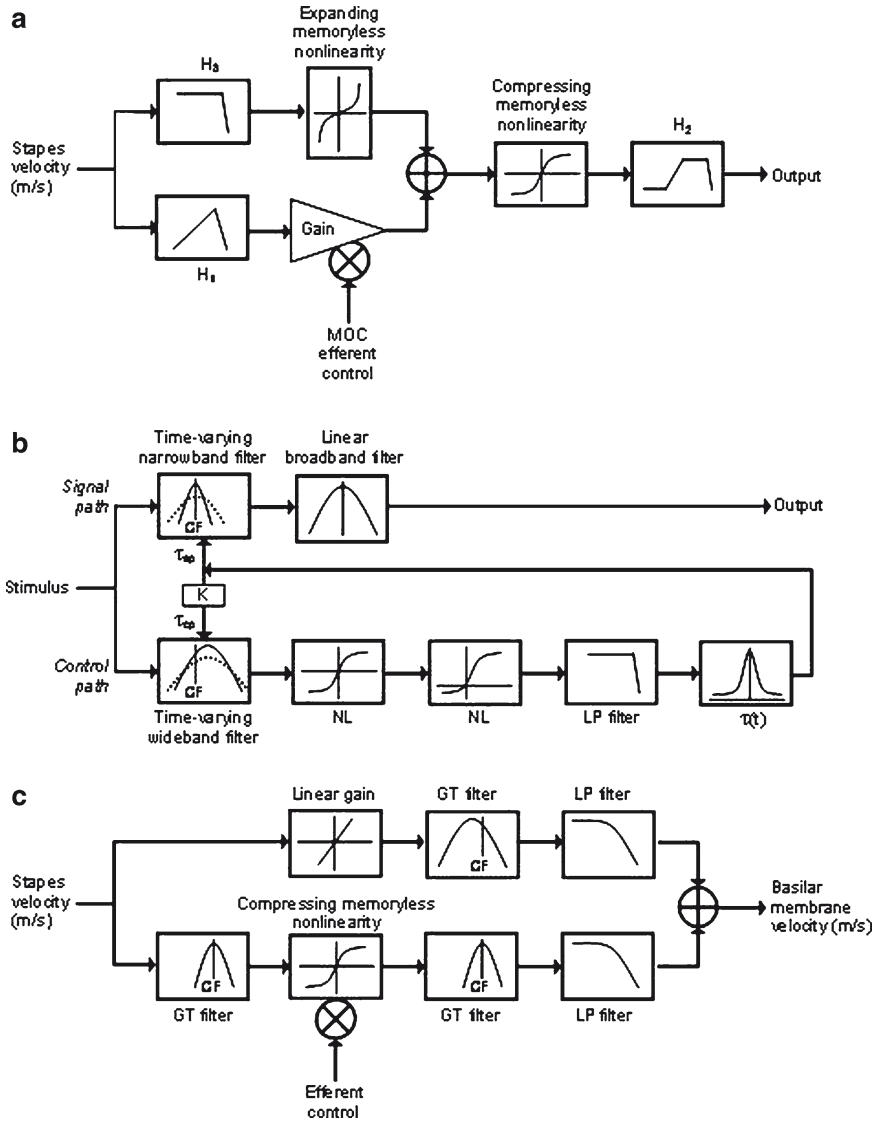

Fig. 2.2 Comparative architecture of three phenomenological nonlinear BM models. (**a**) The multiple bandpass nonlinear filter of Goldstein (adapted from Goldstein 1990). (**b**) The model of Zhang et al. (adapted from Zhang et al. 2001). (**c**) The dual-resonance nonlinear filter of Meddis et al. (adapted from Lopez-Poveda and Meddis 2001). See text for details. *GT* gammatone; *LP* low-pass; *NL* nonlinearity; *MOC* medio-olivocochlear

the bandpass nonlinear filter of Pfeiffer (1970) and Duifhuis (1976). The narrow and broad filters account for BM tuning at low and high levels, respectively. By carefully choosing their shapes and the gain of the compressive gain, the model reproduces level-dependent tuning and BF shifts (Goldstein 1990).

The model was specifically designed to reproduce the nonlinear cyclic interactions between a moderate-level tone at CF and another highly intense tone with a very low frequency, a phenomenon usually referred to as "simple-tone interaction" (or simple-tone interference; Patuzzi et al. 1984). This required incorporating an expanding nonlinearity (inverse in form to the compressing nonlinearity) whose role in the model is to enhance the low frequencies before they interact with on-CF tones at the compressive stage (Fig. 2.2a). With this expanding nonlinearity, the model reproduces detailed aspects of BM suppression and combination tones (Goldstein 1995). However, propagation of combination tones is lacking in the model, although it appears necessary to account for the experimental data regarding the perception of the $2f_1 - f_2$ combination tone (Goldstein 1995).

The MBPNL model was further developed into a version capable of reproducing the response of the whole cochlear partition by means of a bank of interacting MBPNL filters (Goldstein 1993). This newer version gave the model the ability to account for propagating combination tones. However, to date systematic tests have not been reported on this MBPNL filterbank.

2.4.1.2 The Gammatone Filter

It is not possible to understand many of the current signal-processing cochlear models without first understanding the characteristics of their predecessor: the gammatone filter. The gammatone filter was developed to simulate the impulse response of AN fibers as estimated by reverse correlation techniques (Flanagan 1960; de Boer 1975; de Boer and de Jongh 1978; Aertsen and Johannesma 1980). The impulse response of the gammatone filter basically consists of the product of two components: a carrier tone of a frequency equal to the BF of the fiber and a statistical gamma-distribution function that determines the shape of the impulse response envelope. One of the advantages of the gammatone filter is that its digital, time-domain implementation is relatively simple and computationally efficient (Slaney 1993), and for this reason it has been largely used to model both physiological and psychophysical data pertaining to auditory frequency selectivity. It has also been used to simulate the excitation pattern of the whole cochlear partition by approximating the functioning of the BM to that of a bank of parallel gammatone filters with overlapping passbands, a filterbank (e.g., Patterson et al. 1992).

On the other hand, the gammatone filter is linear, thus level independent, and it has a symmetric frequency response. Therefore, it is inadequate to model asymmetric BM responses. Several attempts have been made to design more physiological versions of the gammatone filter. For instance, Lyon (1997) proposed an all-pole digital version of the filter with an asymmetric frequency response. This all-pole version also has the advantage of being simpler than the conventional gammatone filter in terms of

parameters, as its gain at center frequency and its bandwidth are both controlled by a single parameter, namely, the quality factor (Q) of the filter (the quality factor of a filter is defined as the ratio of the filter center frequency, f_C, to the filter bandwidth, BW, measured at a certain number of decibels below the maximum gain, $Q=f_C/\mathrm{BW}$).

2.4.1.3 The Gammachirp Filter

The gammachirp filter of Irino and Patterson (1997), like the all-pole gammatone filter, was designed to produce an asymmetric gammatone-like filter. This was achieved by making the carrier-tone term of the analytic impulse response of the gammatone filter modulated in frequency, thus the suffix chirp. This property was inspired by the fact that the impulse responses of the BM and of AN fibers are also frequency modulated (Recio et al. 1998; Carney et al. 1999).

In its original form, the gammachirp filter was level independent (linear), hence inadequate to simulate the nonlinear, compressive growth of BM response with level. Further refinements of the filter led to a compressive gammachirp filter with a level-independent chirp (Irino and Patterson 2001), hence more consistent with the physiology. The compressive gammachirp filter can be viewed as a cascade of three fundamental filter elements: a gammatone filter followed by a low-pass filter, followed by a high-pass filter with a level-dependent corner frequency. Combined, the first two filters produce an asymmetric gammatone-like filter, which can be approximated to represent the "passive" response of the BM. Because of its asymmetric frequency response, the associated impulse response of this "passive" filter shows a chirp.

The third element in the cascade, the high-pass filter, is responsible for the level dependent gain and tuning characteristics of the compressive gammachirp filter. It is designed to affect only frequencies near the center frequency of the gammatone filter in a level-dependent manner. At low levels, its corner frequency is configured to compensate for the effect of the low-pass filter, thus making the frequency response of the global gammachirp filter symmetric. At high levels, by contrast, its corner frequency is set so that the frequency response of the "passive" filter is almost unaffected and thus asymmetric. The chirping properties of the gammachirp filter are largely determined by those of its "passive" asymmetric filter at all levels, and have been shown to fit well those of AN fibers (Irino and Patterson 2001).

The compressive gammachirp filter has proved adequate to design filterbanks that reproduce psychophysically estimated human auditory filters over a wide range of center frequencies and levels (Patterson et al. 2003). It could probably be used to simulate physiological BM iso-intensity responses directly, although no studies have been reported to date aimed at testing the filter in this regard. Its BF shifts with level as do BM and AN iso-intensity curves, but the trends shown by Irino and Patterson (2001) are not consistent with the physiological data (Tan and Carney 2003). More importantly, we still lack detailed studies aimed at examining the ability of this filter to account for other nonlinear phenomena such as level-dependent

phase responses, combination tones, or two-tone suppression. Some authors have suggested that it cannot reproduce two-tone suppression because it is not a "true" nonlinear filter, but rather a "quasilinear" filter whose shape changes with level (Plack et al. 2002). Recently, a dynamic (time-domain) version of the compressive gammachirp filter adequate for processing time-varying signals has become available (Irino and Patterson 2006).

2.4.1.4 The Model of Carney and Colleagues

Carney and colleagues (Heinz et al. 2001; Zhang et al. 2001) have proposed an improved version of Carney's (1993) composite phenomenological model of the AN response that reproduces a large number of nonlinear AN response characteristics. A version of this model (Tan and Carney 2003) also reproduces level-independent frequency glides (the term "frequency glide" is synonymous with the term "chirp" and both refer to the frequency-modulated character of BM and AN impulse responses).

An important stage of this composite AN model is designed to account for the nonlinear response of a single BM cochlear site (Fig. 2.2b). In essence, it consists of a gammatone filter whose gain and bandwidth vary dynamically in time depending on the level of the input signal (this filter is referred to in the original reports as "the signal path"). For a gammatone filter, both these properties, gain and bandwidth, depend on the filter's time constant, τ (see Eq. (2) of Zhang et al. 2001). In the model, the value of this time constant varies dynamically in time depending on the amplitude of the output signal from a feed-forward control path, which itself depends on the level of the input signal. As the level of the input signal to the control path increases, then the value of τ decreases, thus increasing the filter's bandwidth and decreasing its gain. The structure of the control path is carefully designed to reflect the "active" cochlear process of the corresponding local basilar-membrane site as well as that of neighboring sites. It consists of a cascade of a wideband filter followed by a saturating nonlinearity. This saturating nonlinearity can be understood to represent the transduction properties of outer hair cells and is responsible for the compressive character of the model input/output response. Finally, the bandwidth of the control-path filter also varies dynamically with time, but it is always set to a value greater than that of the signal-path filter. This is necessary to account for two-tone suppression, as it allows for frequency components outside the pass-band of the signal-path filter to reduce its gain and thus the net output amplitude.

This model uses symmetric gammatone filters and, therefore, does not produce asymmetric BM frequency responses or click responses showing frequency glides. The model version of Tan and Carney (2003) solves these shortcomings by using asymmetrical digital filters that are designed in the complex plane (i.e., by positioning their poles and zeros) to have the appropriate glide (or "chirp"). Further, by making the relative position of these poles and zeros in the complex plane independent of level, the model can also account for level-independent frequency glides, consistent with the physiology (de Boer and Nuttall 1997; Recio et al. 1998; Carney et al. 1999).

2.4.1.5 The DRNL Filter of Meddis and Colleagues

The Dual-Resonance NonLinear (DRNL) filter model of Meddis and co-workers (Lopez-Poveda and Meddis 2001; Meddis et al. 2001; Lopez-Poveda 2003) simulates the velocity of vibration of a given site on the BM (Fig. 2.2c). This filter is inspired by Goldstein's MBPNL model and its predecessors (see earlier), although the structure of the DRNL filter is itself unique. The input signal to the filter is processed through two asymmetric bandpass filters arranged in parallel: one linear and broadly tuned, and one nonlinear and narrowly tuned. Gammatone filters are employed that are made asymmetric by filtering their output through a low-pass filter. A compressing memoryless (i.e., instantaneous) gain is applied to the narrow filter that produces linear responses at low levels but compressive responses for moderate levels. The output from the DRNL filter is the sum of the output signals from both paths. Level-dependent tuning is achieved by setting the relative gain of the two filter paths so that the output from the narrow and broad filters dominate the total filter response at low and high levels, respectively. Level-dependent BF shifts are accounted for by setting the center frequency of the broad filter to be different from that of the narrow filter.

The model reproduces suppression because the narrow nonlinear path is actually a cascade of a gammatone filter followed by the compressive nonlinearity, followed by another gammatone filter (Fig. 2.2c). For a two-tone suppression stimulus, the first gammatone filter passes both the suppressor and the probe tone, which are then compressed together by the nonlinear gain. Because the probe tone is compressed with the suppressor, its level at the output of the second filter is less than it would be if it were presented alone. Some versions of the DRNL filter assume that the two gammatone filters in this pathway are identical (Lopez-Poveda and Meddis 2001; Meddis et al. 2001; Sumner et al. 2002), while others (e.g., Plack et al. 2002) allow for the two filters to have different center frequencies and bandwidths to account for suppression phenomena more realistically (specifically, it can be assumed that the first filter is broader and has a higher center frequency than the second filter). On the other hand, the characteristics of the first gammatone filter in this nonlinear pathway determine the range of primary frequencies for which combination tones occur, while the second gammatone filter determines the amplitude of the generated combination tones.

The DRNL filter has proved adequate to reproduce frequency- and level-dependent BM amplitude responses for a wide range of CFs (Meddis et al. 2001; Lopez-Najera et al. 2007). It also reproduces local combination tones (i.e., combination tones that originate at BM regions near the measurement site) and some aspects of two-tone suppression (Meddis et al. 2001; Plack et al. 2002). Its impulse response resembles that of the BM and it shows frequency glides (Meddis et al. 2001; Lopez-Najera et al. 2005). These characteristics, however, appear very sensitive to the values of the model parameters, particularly to the total order of the filters in both paths and to the frequency response of the middle-ear filter used in the model (Lopez-Najera et al. 2005).

Filterbank versions of the DRNL filter have been proposed for human (Lopez-Poveda and Meddis 2001), guinea pig (Sumner et al. 2003b), and chinchilla (Lopez-Najera et al. 2007) based on corresponding experimental data. These filterbanks

do not consider interaction between neighboring filters or propagation of combination tones. The parameters of the DRNL filter may be simply adjusted to model BM responses in cochleae with OHC loss (Lopez-Poveda and Meddis 2001). A version of the DRNL exists designed to account for effect of efferent activation on BM responses (Ferry and Meddis 2007).

This filter has been successfully employed for predicting the AN representation of stimuli with complex spectra, such as HRTF (Lopez-Poveda 1996), speech (Holmes et al. 2004), harmonic complexes (Gockel et al. 2003; Wiegrebe and Meddis 2004), or amplitude-modulated stimuli (Meddis et al. 2002). The model has also been used to drive models of brain stem units (Wiegrebe and Meddis 2004). It has also been used as the basis to build a biologically inspired speech processor for cochlear implants (Wilson et al. 2005, 2006; see also Chapter 9).

2.5 Inner Hair Cells

IHCs are responsible for the mechanoelectrical transduction in the organ of Corti of the mammalian cochlea. Deflection of their stereocilia toward the tallest cilium in the bundle increases the inward flow of ions and thus depolarizes the cell. Stereocilia deflection in the opposite direction closes transducer channels and prevents the inward flow of ions to the cell. This asymmetric gating of transducer channels has led to the well-known description of the IHC as a half-wave rectifier. Potassium (K^+) is the major carrier of the transducer current. The "excess" of intracellular potassium that may result from bundle deflections is eliminated through K^+ channels found in the IHC basolateral membrane, whose conductance depends on the IHC basolateral transmembrane potential (Kros and Crawford 1990). Therefore, the intracellular voltage variations produced by transducer currents may be modulated also by currents flowing through these voltage-dependent basolateral K^+ conductances. The intracellular voltage is further determined by the capacitive effect of the IHC membrane and by the homeostasis of the organ of Corti.

The in vivo IHC inherent input/output response characteristics are hard to assess because in vivo measurements reflect a complex combination of the response characteristics of the middle ear, the BM, and the IHC itself (Cheatham and Dallos 2001). Inherent IHC input/output functions have been inferred from measurements of the growth of the AC or DC components of the receptor potential with increasing sound level for stimulus frequencies an octave or more below the characteristic frequency of the IHC. The BM responds linearly to these frequencies (at least in basal regions). Therefore, any sign of nonlinearity is attributed to inherent IHC processing characteristics (Patuzzi and Sellick 1983). These measurements show that the dc component of the receptor potential grows expansively (slope of 2 dB/dB) with increasing sound level for sound levels near threshold and that the AC and DC components of the receptor potential grow compressively (slope <1 dB/dB) for moderate to high sound levels (Patuzzi and Sellick 1983). These nonlinear transfer characteristics reflect the combination of nonlinear activation of

transducer and basolateral K^+ currents (described by Lopez-Poveda and Eustaquio-Martín 2006).

The in vivo IHC inherent frequency response is also difficult to assess (Cheatham and Dallos 2001). Some authors have estimated it as the ratio of the AC to the DC components of the in vivo receptor potential (AC/DC ratio) on the assumption that this ratio is normalized for constant input to the cell (Sellick and Russell 1980). The AC/DC ratio decreases with increasing stimulus frequency (Russel and Sellick 1978). This low-pass filter effect is attributed to the resistor-capacitance properties of the IHC membrane. To a first approximation, this is independent of the driving force to the cell (Russel and Sellick 1978) and of the cell's membrane potential (cf. Kros and Crawford 1990; Lopez-Poveda and Eustaquio-Martín 2006). Therefore, it is considered that the low-pass filter behavior is independent of sound level (Russel and Sellick 1978). This low-pass filter effect is thought to be responsible for the rapid roll-off of AN phase-locking with increasing frequency above approximately 1.5–2 kHz (Palmer and Russell 1986) and has led to the common description of the IHC as a low-pass filter.

It is worth mentioning that while the AC/DC ratio shows a low-pass frequency response, the AC component alone shows a bandpass response tuned at a frequency of approximately 500 Hz (Sellick and Russell 1980) or 1 kHz (Dallos 1984, 1985) for low sound levels. This result is important because it is for a basal IHC in response to low-frequency stimuli. The excitation of basal BM sites is linear and untuned in response to low-frequency tones. Therefore, the result of Sellick and Russell (1980) constitutes direct evidence for bandpass AC responses without substantial contributions from BM tuning. They argued that the rising slope of the response indicates that the IHC receptor potential responds to BM velocity for frequencies below approximately 200 Hz and to BM displacement above that frequency (see also Shamma et al. 1986).

The IHC responds nonlinearly also in time. The time-dependent activation of basolateral K^+ channels induces a nonlinear, time-dependent adaptation of the receptor potential (Kros and Crawford 1990) that could contribute to adaptation as observed in the AN (Kros 1996). This in vitro result, however, is awaiting confirmation in vivo, but computational modeling studies support this suggestion (Zeddies and Siegel 2004; Lopez-Poveda and Eustaquio-Martín 2006).

2.5.1 Approaches to Modeling the IHC Transfer Function

IHC models aim to simulate the cell's intracellular potential in response to BM excitation because the latter determines the release of neurotransmitter from within the IHC to the synaptic cleft. It is common to model the function of the IHC using either biophysical analogs or signal-processing analogs. The latter consider the IHC as a cascade of an asymmetric, saturating nonlinear gain, which accounts for the activation of the transducer currents, followed by a low-pass filter, which accounts for the resistor-capacitor filtering of the IHC membrane. The order and cutoff frequency

of this filter are chosen so as to mimic as closely as possible the physiological low-pass characteristics of the IHC.

These signal-processing models are easy to implement, fast to evaluate, and require very few parameters. For these reasons, they are widely used in composite peripheral auditory models (e.g., Robert and Eriksson 1999; Zhang et al. 2001). However, they neglect important aspects of IHC processing and are limited in scope. For instance, IHCs are modeled as a low-pass filter regardless of whether the input to the IHC model stage is BM velocity or displacement. As discussed in the preceding section, this is almost certainly inappropriate for sounds with frequencies below 0.2–1 kHz. In addition, these models do not account for the time-activation of basolateral K^+ currents, which could be significant, particularly for brief and intense sounds (Kros 1996). Another shortcoming is that their parameters do not represent physiological variables; hence they do not allow modeling some forms of hearing loss associated to IHC function without changing the actual transducer and/or filter function (see Chapter 7).

An alternative approach is to model the IHC using biophysical models (an early review is provided by Mountain and Hubbard 1996). Typically these are electrical-circuit analogs of the full organ of Corti. The model of Lopez-Poveda and Eustaquio-Martín (2006) is an example. It consists of several elements that describe the electrical properties of the apical and basal portions of the IHC and its surrounding fluids. The model assumes that the intracellular space is equipotential and thus can be represented by a single node. It assumes that the IHC intracellular potential is primarily controlled by the interplay of a transducer, variable (inward) K^+ current that results from stereocilia deflections and a basolateral (outward) K^+ current that eliminates the excess of intracellular K^+ from within the IHC. The magnitude of the transducer current is calculated from stereocilia displacement using a Boltzmann function that describes the gating of transducer channels. The excess of intracellular K^+ is eliminated through two voltage- and time-dependent nonlinear activating basolateral conductances, one with fast and one with slow-activation kinetics. The activation of these two conductances is modeled using a Hodgkin–Huxley approach. The reversal potential of each of the currents involved is accounted for by a shunt battery. The capacitive effects of the IHC membrane are modeled with a single capacitor. The flow of transducer current depends also on the endocochlear potential, which is simulated with a battery.

This relatively simple electrical circuit accounts for a wide range of well reported in vitro and in vivo IHC response characteristics without a need for readjusting its parameters across data sets. Model simulations support the idea that the basolateral K^+ conductances effectively reduce the rate of growth of IHC potential with increasing stereocilia displacement by more than a factor of two for displacements above approximately 5 nm. Such compression affects the DC component of the cell's potential in a similar way for all stimulation frequencies. The AC component is equally affected but only for stimulation frequencies below 800 Hz. The simulations further suggest that the nonlinear gating of the transducer current produces an expansive growth of the DC potential with increasing sound level (slope of 2 dB/dB) at low sound pressure levels (Lopez-Poveda and Eustaquio-Martín 2006).

The model of Shamma et al. (1986) is similar and simpler in that it considers voltage- and time-independent basolateral K^+ currents. A more sophisticated version of the model of Lopez-Poveda and Eustaquio-Martín (2006) exists that incorporates the role of transmembrane cloring and sodium currents and pumps in shaping the IHC intracellular potential (Zeddies and Siegel 2004).

Biophysical IHC models have been used successfully in composite models of the peripheral auditory system (e.g., Sumner et al. 2002, 2003a, b). In these cases, a high-pass filter is used to couple BM displacement to stereocilia displacement.

2.6 Auditory Nerve Synapse

AN activity is provoked by the release of transmitter substance (glutamate) into the synaptic cleft between the AN dendrites and the IHC. The rate of release of this transmitter is regulated by two factors, the IHC receptor potential and the availability of transmitter in the presynaptic area. These two processes can be modeled separately.

Researchers generally agree that vesicles of transmitter substance are held inside the cell in a local store close to the synaptic site from which the vesicles are released into the postsynaptic cleft between the cell and a dendrite of an AN fiber. As the electrical potential inside the cell increases, the probability of release of one or more vesicles also increases. The number of vesicles available for release is relatively small and a series of release events will result in a depletion of the available vesicle store. When this happens, the rate of release of vesicles falls even though the receptor potential is unchanged. The rate will remain depressed until the presynaptic store can be replenished (Smith and Zwislocki 1975; Smith et al. 1985). It is important to distinguish between the probability that a vesicle will be released (if it is available) and the number of vesicles available for release. The vesicle release rate is the product of these two values. If no transmitter is available for release, then none will be released even if the probability of release is high. In Fig. 2.1, the "release probability" in the second from bottom panel is the first of these two quantities.

The reduction of AN spike rate after stimulation is known as "adaptation." The speed of recovery from adaptation is thought to reflect the rate at which the available store can be replenished. While there is considerable uncertainty concerning the details of this process, it nevertheless remains an important goal for the modeler to generate an accurate representation of this process. This is because it is reflected in many aspects of psychophysics where sounds are presented in rapid succession, each influencing the response of later sounds as a function of the resulting depletion of the available pool of transmitter vesicles.

2.6.1 Calcium Control of Transmitter Release

Most early models of the transmitter release and recovery proposed a simple relationship between the receptor potential level and rate of release of transmitter

(Siebert 1965; Weiss 1966; Eggermont 1973; Schroeder and Hall 1974; Oono and Sujaku 1975; Nilsson 1975; Geisler et al. 1979; Ross 1982; Schwid and Geisler 1982; Smith and Brachman 1982). In so doing, they ignored the complex nature of the relationship. This was because research has only recently unraveled the details (see, e.g., Augustine et al. 1985). It is now known that the release of transmitter is only indirectly controlled by the internal voltage of the cell. Instead, the voltage controls the rate of flow of calcium into the cell and it is this calcium that promotes the release of available transmitter into the synaptic cleft.

While it might be thought that this is one complication too many, there are indications that it is an essential part of an understanding of the signal processing that occurs at this stage. For example, Kidd and Weiss (1990) have suggested that delays associated with the movement of calcium contribute to the reduction of AN phase-locking at high frequencies. Phase-locking is already limited by the IHC membrane capacitance (see earlier) but they suggest that the rate of accumulation of presynaptic calcium further limits this effect. To some extent this is inevitable and much depends on an exact knowledge of the rate of accumulation.

More recently, it has been suggested that the accumulation of presynaptic calcium might be the physiological basis for some aspects of psychophysical thresholds (Heil and Neubauer 2003). Sumner et al. (2003a) and Meddis (2006) have also suggested that differences in the rate of accumulation and dissipation of calcium might control the rate/level function of the fiber attached to the synapse, particularly the difference between low and high spontaneous rate (LSR, HSR) fibers. The synapse is very inaccessible and difficult to study. As a consequence, these ideas must remain speculative but they do justify the inclusion of the calcium control stage in recent models of transmitter release.

Calcium enters the cell through voltage-gated calcium ion channels located close to the synapse. The number of open calcium channels is determined by the receptor potential; as the voltage rises, more gates open. Calcium ions enter the cell and accumulate in the region of the synapse. The density of ions close to the synapse determines the probability that a transmitter vesicle will be released into the cleft. However, the calcium dissipates rapidly or is chemically inactivated by a process known as buffering and the calcium concentration falls rapidly if the receptor potential falls again. The opening and closing of these ion channels as well as calcium accumulation and dissipation can be modeled using equations that are generally agreed upon among physiologists (Meddis 2006).

2.6.2 Transmitter Release

Transmitter release is an important feature of auditory models because it is the basis for explaining adaptation in the AN. From the beginning, all models of the auditory periphery have included a stage that simulates this process of depletion and recovery. All assume that there is a reservoir of transmitter that releases its contents into the synaptic cleft at a rate proportional to the stimulus intensity.

Although this is a satisfactory model for many purposes, the data suggest that the situation is more complex. If only one reservoir is involved, we might expect only one time constant of adaptation when a stimulus is presented. However, the data indicate two or even three time constants (Smith and Brachman 1982). The same applies to the recovery process where the time course of recovery is complex (Harris and Dallos 1979). The most elegant solution to this problem was proposed by Westerman and Smith (1984, 1988), who suggested a cascade of reservoirs each with their own time constant (Fig. 2.3). When the reservoir closest to the synapse becomes depleted, it is slowly refilled by the reservoir immediately above it. The third reservoir refills the second and so on. In a cascade system, the time constants of all three reservoirs are reflected in the time course of release of transmitter from the pre-synaptic reservoir. Westerman's ideas have been adopted in the modeling of Carney (1993).

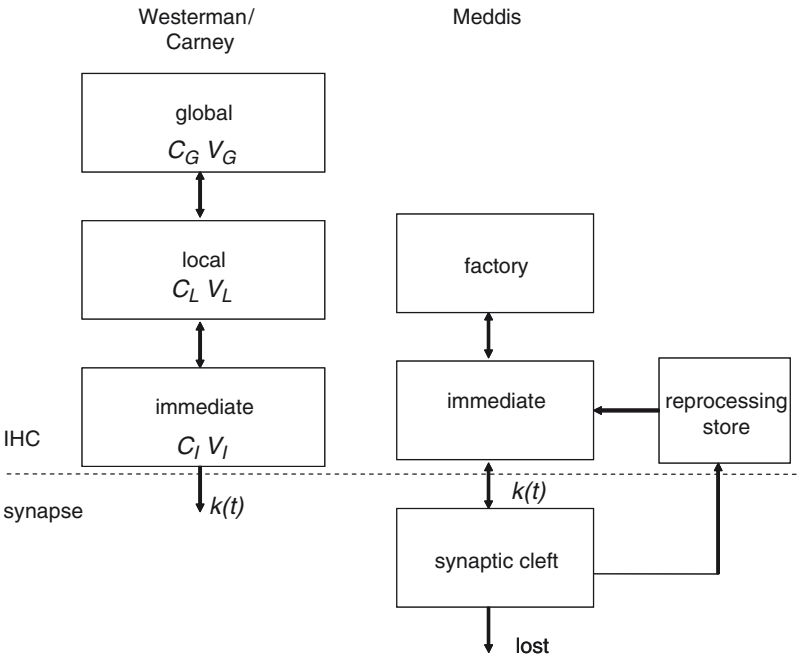

Fig. 2.3 Westerman/Carney and Meddis models of IHC/AN transmitter release. In both models $k(t)$ represents the rate at which transmitter substance is released into the synaptic cleft and this is indirectly controlled by the receptor potential of the IHC. In the Westerman/Carney model, C represents the concentration of transmitter in a reservoir and V represents its volume. P is the permeability of the path between two reservoirs. The *dashed line* indicates the IHC membrane that the transmitter must cross when released into the cleft. Equations controlling the model can be found in Zhang and Carney (2005). The Meddis model consists of reservoirs containing individual vesicles of transmitter (usually less than 20 vesicles). The equations controlling the probability that one vesicle is transferred from one reservoir to another can be found in Meddis (2006). The two models are arranged slightly differently but the behavior of the two systems is very similar

Meddis (1986, 1988) suggested an alternative system that also involved reservoirs of transmitter but used reuptake of transmitter from the synaptic cleft as the major source of replenishment of the presynaptic reservoir. Zhang and Carney (2005) have recently reevaluated both models and found that they are mathematically very similar. Recent studies of IHC physiology have confirmed that reuptake of transmitter does take place but on a much longer time scale than required by the Meddis model (see Griesinger et al. 2002).

Models of transmitter circulation are relatively straightforward and consist of a cascade of reservoirs with transmitter flowing between them. The flow of transmitter between reservoirs is determined by the relative concentrations of transmitter in the reservoirs as well as the permeability of the channels linking them. Details of the equations used to evaluate both models can be found in Zhang and Carney (2005) and Meddis (2006). The two models are illustrated in Fig. 2.3.

The most important reservoir is the "immediate" pool that releases transmitter into the synaptic cleft according to the level of the receptor potential. After stimulation, this pool becomes depleted and fewer vesicles are available for release, leading to adaptation of the response. It is important to note that the receptor potential is not affected during adaptation. The reduction in transmitter release is accounted for mainly by the reduction in available transmitter. Recovery takes place over time and as the result of replenishment either from transmitter reuptake (Meddis 1988) or a from a "global" reserve reservoir (Westerman and Smith 1988; Carney 1993).

2.7 Auditory Nerve Activity

The release of transmitter is generally agreed to be a stochastic process. The instantaneous probability of release is determined by the product of the concentration of presynaptic calcium and the number of available transmitter vesicles. However, the release event is itself a random outcome. Stochastic release of transmitter can be generated simply using random number generators to convert the release probabilities into binary release events. It is not known exactly how release events translate into AN spike events. Meddis (2006) makes the simplifying assumption that a single vesicle release event is enough to trigger an AN spike. This idea was based on some early observations of postsynaptic events by Siegel (1992). Goutman and Glowatzki (2007) offer some recent support for this view but the issue is the focus of continuing research. Certainly, the assumption of the model works well in practice.

Modelers often use the release rate as the final result of the modeling exercise. In the long run, the rate of release is a useful indication of the rate of firing of the AN fiber attached to the synapse. This is a quick and convenient representation if the model is to be used as the input to another computationally intensive application such as an automatic speech recognition device.

Modeling individual spike events in AN fibers is more time-consuming than computing probabilities alone but for many purposes it is essential, for example, when the next stage in the model consists of models of neurons in the brain stem.

Refractory effects should be included in the computation for greater accuracy. In common with other nerve cells, the AN fiber is limited in terms of how soon it can fire immediately after a previous spike. There is an absolute limit (~500 ms) on how soon a second spike can occur. The absolute refractory period is followed by a relative refractory period during which time the probability of an action potential recovers exponentially. Carney (1993) describes a useful method to simulate such effects.

2.8 Efferent Effects

So far we have considered the auditory periphery in terms of a one-way path, from the eardrum to the AN. In reality, many fibers travel in the other direction from the brain stem to the cochlea. Efferent feedback operates through two separate systems: lateral and medial (Guinan 2006). The lateral system acts directly on the dendrites of afferent auditory nerve fibers and is only poorly understood. The medial system acts by damping the response of the BM indirectly through the OHCs. This damping effect modifies the relationship between the stimulus level and the BM response. This reduced response also leads to less adaptation in the auditory nerve. It is widely believed that this latter effect is critical to the function of the medial efferent system by protecting the periphery from overstimulation.

The function of these efferent fibers is largely unknown and they rarely feature in computer models. A computer model has been developed (Ghitza et al. 2007; Messing et al. 2009) showing that efferent feedback can improve vowel discrimination against a background of noise. Ferry and Meddis (2007) have also shown that a model with efferent feedback can simulate physiological observations at the level of the BM and the AN.

2.9 Summary

It can be seen that a model of the auditory periphery is very complex. It is composed of many stages, each of which has its own associated scientific literature. Individual component stages are always compromises in terms of simulation accuracy. Part of the problem is the need to compute the result in a reasonable amount of time but it is also the case that researchers have not yet finally agreed on the details of any one processing stage. Models will need to change as new data and new insights are published. Nevertheless, models are already good enough to use them in a range of applications.

The nonlinear nature of the auditory periphery has many unexpected consequences, and it is important that the user of any model should appreciate from the outset that a computer model of the auditory periphery is not simply a biological way to generate a spectral analysis of the input sound. The ear appears to be doing something quite different. Figure 2.4 gives a simple example of a nonlinear effect that would not

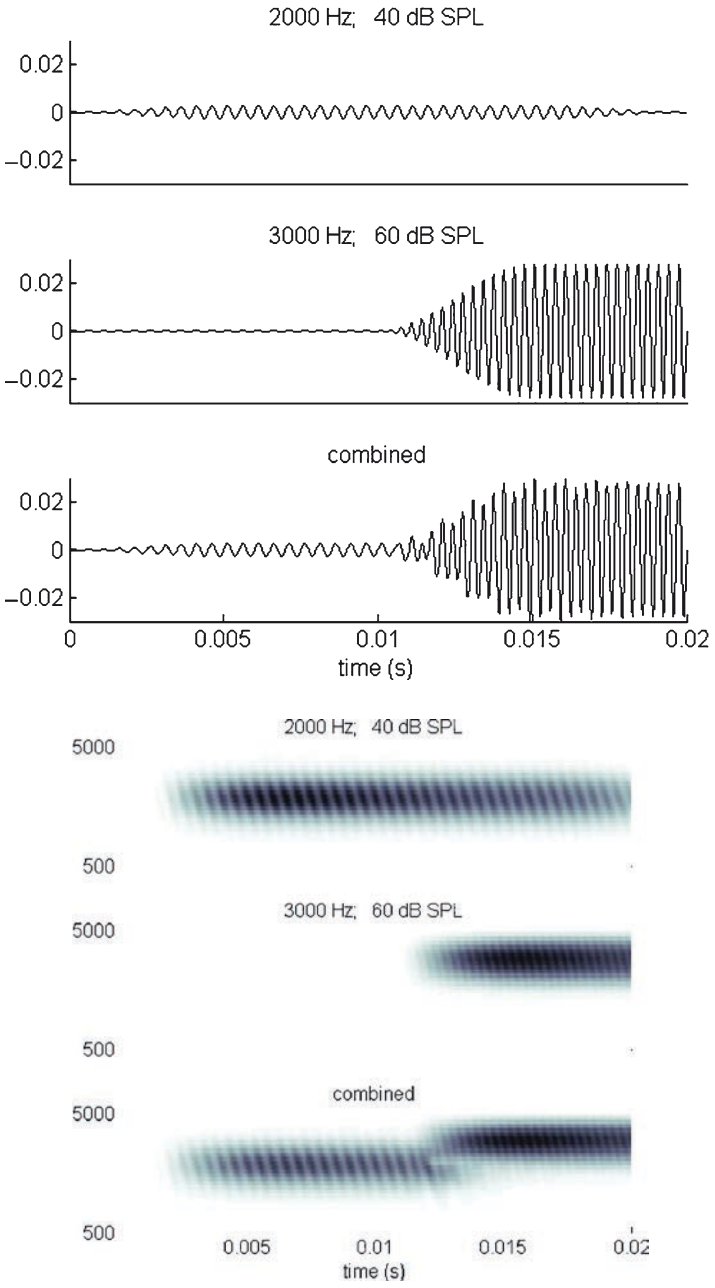

Fig. 2.4 A demonstration of two-tone suppression in a computer model of the auditory periphery. The model uses 30 channels with best frequencies distributed between 500 and 5 kHz. *Left*: Stimuli, all presented on the same scale. *Right*: Multichannel model showing probability of transmitter release. *Top panels*: 2-kHz, 20-ms tone (the probe) presented at 40 dB SPL. *Middle panels*: 3-kHz, 10-ms tone (the suppressor) presented at 60 dB SPL. *Bottom panels*: both tones presented together. The response to the probe tone is reduced when the suppressor begins

be seen in a discrete Fourier transform. The top panel shows the response to a single pure tone called the "probe." The second panel shows the response to a second pure tone called the "suppressor." Note that the suppressor is timed to start after the probe. The third panel shows what happens when the two tones are combined. When the suppressor tone starts, the response to the probe is substantially reduced. This is a consequence of the nonlinearities in the model and would never occur in a linear system. While this demonstration is very clear, it should not be assumed that all tones suppress all other tones. This effect occurs only with certain combinations of levels and tone frequencies. This example was found only after careful searching for an ideal combination.

Another difference from traditional signal processing can be seen with background firing rates in the auditory nerve. The majority of auditory nerve fibers are spontaneously active. They have spontaneous firing rates up to 100 spikes/s. When the fiber is driven by a steady high intensity tone, its firing rate will rarely exceed 300 spikes/s. Figure 2.5 shows the response of an auditory model to speech (the utterance "one-oh seven") at three speech levels. Two kinds of output are shown. The left-hand panels show the pattern of transmitter release rates while the right-hand panels show raster plots of spike activity in a single fiber per channel. Release rates are faster to compute and show a much clearer picture. The spiking activity is much less easy to interpret, but it must be remembered that a full model has thousands of

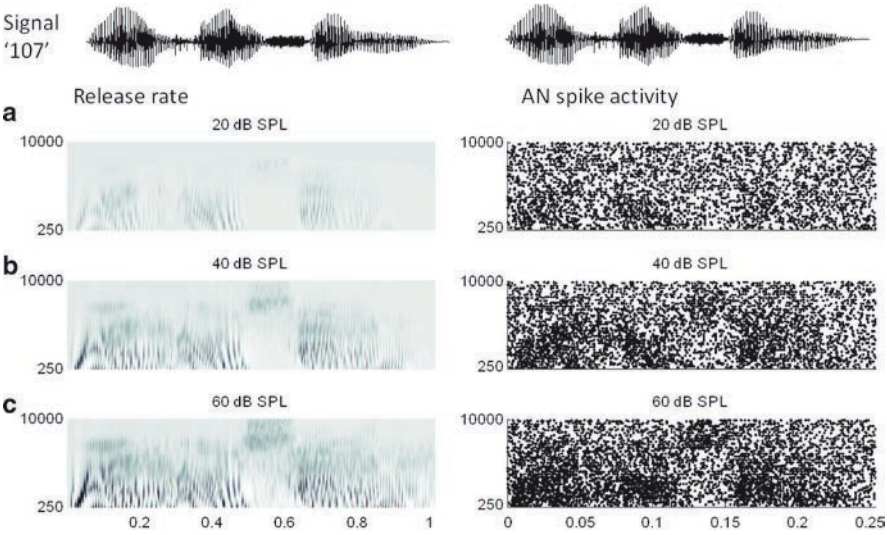

Fig. 2.5 Sixty-channel model AN response to the speech utterance "one oh seven" presented at three signal levels 20, 40, and 60 dB SPL. Channel best frequencies ranged between 250 Hz and 10 kHz. The model is based on equations in Meddis (2006). *top*: Transmitter vesicle release rate. *bottom*: Raster plot of individual AN fibers (1 per channel). The conventions used here are also explained in Fig. 2.1

fibers and the aggregate activity of all the fibers will follow the release rate pattern very closely (except for the refractory effects that are built into the fiber activity but not the transmitter release rates). The release rates are easier to interpret and link to the input signal but the spiking activity is shown to remind the reader that this is the true output of the model. This is what will be passed to later processing modules representing the activity in the cochlear nucleus. Clearly, the background activity of the fibers and the stochastic nature of the response present important challenges to the signal processing power of the brain stem neurons that receive AN input.

References

Aertsen AM, Johannesma PI (1980) Spectro-temporal receptive fields of auditory neurons in the grassfrog: I. Characterization of tonal and natural stimuli. Biol Cybern 38:223–234.

Aibara R, Welsch JT, Puria S, Goode RL (2001) Human middle-ear transfer function and cochlear input impedance. Hear Res 152:100–109.

Algazi VR, Duda RO, Morrison RP, Thompson DM (2001) Structural composition and decomposition of HRTFs. In: Proceedings of 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. New Paltz, NY, pp. 103–106.

Augustine GJ, Charlton MP, Smith SJ (1985) Calcium entry into voltage-clamped pre-synaptic terminals of squid. J Physiol 367:143–162.

Bacon SP (2004) Overview of auditory compression. In: Bacon SP, Fay RR, Popper AN (eds), Compression: From Cochlea to Cochlear Implants. New York: Springer, pp. 1–17.

Burkhard MD, Sachs RM (1975) Anthropometric manikin for acoustic research. J Acoust Soc Am 58:214–222.

Carlile S, Pralong D (1994) The location-dependent nature of perceptually salient features of the human head-related transfer functions. J Acoust Soc Am 95:3445–3459.

Carlile S, Martin R, McAnally K (2005) Spectral information in sound localization. Int Rev Neurobiol 7:399–435.

Carney LH (1993) A model for the responses of low-frequency auditory-nerve fibers in cat. J Acoust Soc Am 93:402–417.

Carney LH, McDuffy MJ, Shekhter I (1999) Frequency glides in the impulse responses of auditory-nerve fibers. J Acoust Soc Am 105:2384–2391.

Cheatham MA, Dallos P (2001) Inner hair cell response patterns: implications for low-frequency hearing. J Acoust Soc Am 110:2034–2044.

Cooper NP (1996) Two-tone suppression in cochlear mechanics. J Acoust Soc Am 99:3087–3098.

Cooper NP (2004) Compression in the peripheral auditory system. In: Bacon SP, Fay RR, Popper AN (eds), Compression: From Cochlea to Cochlear Implants. New York: Springer, pp. 19–61.

Dallos P (1984) Some electrical circuit properties of the organ of Corti: II. Analysis including reactive elements. Hear Res 14:281–291.

Dallos P (1985) Response characteristics of mammalian cochlear hair cells. J Neurosci 5: 1591–1608.

de Boer E (1975) Synthetic whole-nerve action potentials for the cat. J Acoust Soc Am 58: 1030–1045.

de Boer E (1996) Mechanics of the cochlea: modeling efforts. In: Dallos P, Popper AN, Fay RR (eds), Auditory Computation. New York: Springer, pp. 258–317.

de Boer E, de Jongh HR (1978) On cochlear encoding: potentialities and limitations of the reverse correlation technique. J Acoust Soc Am 63:115–135.

de Boer E, Nuttall AL (1997) The mechanical waveform of the basilar membrane: I: Frequency modulation (“glides”) in impulse responses and cross-correlation functions. J Acoust Soc Am 101:3583–3592.

Deng L, Geisler CD (1987) A composite auditory model for processing speech sounds. J Acoust Soc Am 82:2001–2012.
Derleth RP, Dau T, Kollmeier B (2001) Modeling temporal and compressive properties of the normal and impaired auditory system. Hear Res 159:132–149.
Dornhoffer JL (1998) Hearing results with the Dornhoffer ossicular replacement prostheses. Laryngoscope 108:531–536.
Duda RO, Martens WL (1998) Range dependence of the response of a spherical head model. J Acoust Soc Am 104:3048–3058.
Duifhuis H (1976) Cochlear nonlinearity and second filter: possible mechanism and implications. J Acoust Soc Am 59:408–423.
Duifhuis H (2004) Comments on "An approximate transfer function for the dual-resonance nonlinear filter model of auditory frequency selectivity." J Acoust Soc Am 115(5 Pt 1):1889–1990.
Eggermont JJ (1973) Analogue modeling of cochlea adaptation. Kybernetic 14:117–126.
Ferry RT, Meddis R (2007) A computer model of medial efferent suppression in the mammalian auditory system. J Acoust Soc Am 122:3519–3526.
Flanagan JL (1960) Models for approximating basilar membrane displacement. Bell Syst Technol J 39:1163–1191.
Gan RZ, Sun Q, Dyer RK, Chang K-H, Dormer KJ (2002) Three-dimensional modeling of middle ear biomechanics and its applications. Otol Neurotol 23:271–280.
Geisler CD, Le S, Schwid H (1979) Further studies on the Schroeder-hall hair-cell model. J Acoust Soc Am 65:985–990.
Ghitza O, Messing D, Delhorne L (2007) Towards predicting consonant confusions of degraded speech. In: Kollmeier B, Klump, G, Hohmann V, Langemann U, Mauermann M, Uppenkamp S, Verhey J (eds), Hearing: From Sensory Processing to Perception. New York: Springer, pp. 541–550.
Gockel H, Moore BCJ, Patterson RD, Meddis R (2003) Louder sounds can produce less forward masking effects: effects of component phase in complex tones. J Acoust Soc Am 114: 978–990.
Goldstein JL (1966) Auditory nonlinearity. J Acoust Soc Am 41:676–689.
Goldstein JL (1988) Updating cochlear driven models of auditory perception: a new model for nonlinear auditory frequency analysing filters. In: Elsendoorn BAG, Bouma H (eds), Working Models of Human Perception. London: Academic, pp. 19–58.
Goldstein JL (1990) Modeling rapid waveform compression on the basilar membrane as multiple-bandpass-nonlinearity filtering. Hear Res 49:39–60.
Goldstein JL (1993) Exploring new principles of cochlear operation: bandpass filtering by the organ of Corti and additive amplification by the basilar membrane. In: Duifhuis H, Horst JW, van Dijk P, van Netten SM (eds), Biophysics of Hair Cell Sensory Systems. Singapore: World Scientific, pp. 315–322.
Goldstein JL (1995) Relations among compression, suppression, and combination tones in mechanical responses of the basilar membrane: data and MBPNL model. Hear Res 89:52–68.
Goode RL, Killion M, Nakamura K, Nishihara S (1994) New knowledge about the function of the human middle ear: development of an improved analog model. Am J Otol 15:145–154.
Goutman JD, Glowatzki E (2007) Time course and calcium dependence of transmitter release at a single ribbon synapse. Proc Natl Acad Sci U S A 104:16341–16346.
Griesinger CB, Richards CD, Ashmore JF (2002) FM1-43 reveals membrane recycling in adult inner hair cells of the mammalian cochlea. J Neurosci 22:3939–3952.
Guinan JJ (2006) Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. Ear Hear 27:589–607.
Guinan JJ, Peake WT (1966) Middle-ear characteristics of anaesthetized cats. J Acoust Soc Am 41:1237–1261.
Harris DM, Dallos P (1979) Forward masking of auditory nerve fiber responses. J Neurophysiol 42:1083–1107.
Heil P, Neubauer H (2003) Unifying basis of auditory thresholds based on temporal summation. Proc Natl Acad Sci U S A 100:6151 6156.

Heinz MG, Zhang X, Bruce IC, Carney LH (2001) Auditory nerve model for predicting performance limits of normal and impaired listeners. Acoust Res Lett Online 2:91–96.
Holmes SD, Sumner CJ, O'Mard LPO, Meddis R (2004) The temporal representation of speech in a nonlinear model of the guinea pig cochlea. J Acoust Soc Am 116:3534–3545.
Irino T, Patterson RD (1997) A time-domain, level-dependent auditory filter: the gammachirp. J Acoust Soc Am 101:412–419.
Irino T, Patterson RD (2001) A compressive gammachirp auditory filter for both physiological and psychophysical data. J Acoust Soc Am 109:2008–2022.
Irino T, Patterson RD (2006) A dynamic, compressive gammachirp auditory filterbank. IEEE Audio Speech Lang Process 14:2222–2232.
Kidd RC, Weiss TF (1990) Mechanisms that degrade and timing information in the cochlea. Hear Res 49:181–208.
Kistler DJ, Wightman FL (1992) A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. J Acoust Soc Am 91:1637–1647.
Kleinschmidt M, Tchorz J, Kollmeier B (1999) Combining speech enhancement and auditory feature extraction for robust speech recognition. Speech Commun 34:75–91.
Koike T, Wada H, Kobayashi T (2002) Modeling of the human middle ear using the finite-element method. J Acoust Soc Am 111:1306–1317.
Kringlebotn M (1988) Network model for the human middle ear. Scand Audiol 17:75–85.
Kros CJ (1996) Physiology of mammalian cochlear hair cells. In: Dallos P, Popper AN, Fay RR (eds), The Cochlea. New York: Springer, pp. 318–385.
Kros CJ, Crawford AC (1990) Potassium currents in inner hair cells isolated from the guinea-pig cochlea. J Physiol 421:263–291.
Kulkarni A, Colburn HS (2004) Infinite-impulse-response models of the head-related transfer function. J Acoust Soc Am 115:1714–1728.
Kulkarni A, Isabelle SK, Colburn HS (1999) Sensitivity of human subjects to head-related-transfer-function phase spectra. J Acoust Soc Am 105:2821–2840.
Lopez-Najera A, Meddis R, Lopez-Poveda EA (2005) A computational algorithm for computing non-linear auditory frequency selectivity: further studies. In: Pressnitzer, D, de Cheveigné A, McAdams S, Collet L (eds), Auditory Signal Processing: Physiology, Psychoacoustics, and Models. New York: Springer, pp. 14–20.
Lopez-Najera A, Lopez-Poveda EA, Meddis R (2007) Further studies on the dual-resonance nonlinear filter model of cochlear frequency selectivity: responses to tones. J Acoust Soc Am 122:2124–2134.
Lopez-Poveda EA (1996) The physical origin and physiological coding of pinna-based spectral cues. PhD thesis, Loughborough University, UK.
Lopez-Poveda EA (2003) An approximate transfer function for the dual-resonance nonlinear filter model of auditory frequency selectivity. J Acoust Soc Am 114:2112–2117.
Lopez-Poveda EA (2005) Spectral processing by the peripheral auditory system: facts and models. Int Rev Neurobiol 70:7–48.
Lopez-Poveda EA, Eustaquio-Martín A (2006) A biophysical model of the inner hair cell: the contribution of potassium current to peripheral compression. J Assoc Res Otolaryngol 7:218–235.
Lopez-Poveda EA, Meddis R (1996) A physical model of sound diffraction and reflections in the human concha. J Acoust Soc Am 100:3248–3259.
Lopez-Poveda EA, Meddis R (2001) A human nonlinear cochlear filterbank. J Acoust Soc Am 10:3107–3118.
Lopez-Poveda EA, Plack CJ, Meddis R (2003) Cochlear nonlinearity between 500 and 8000 Hz in normal-hearing listeners. J Acoust Soc Am 113:951–960.
Lopez-Poveda EA, Barrios LF, Alves-Pinto A (2007) Psychophysical estimates of level-dependent best-frequency shifts in the apical region of the human basilar membrane. J Acoust Soc Am 121:3646–3654.
Lyon RF (1997) All-pole models of auditory filtering. In: Lewis ER, Lyon R, Long GR, Narins PM (eds), Diversity in Auditory Mechanics. Singapore: World Scientific, pp. 205–211.

Meddis R (1986) Simulation of mechanical to neural transduction in the auditory receptor. J Acoust Soc Am 79:702–711.
Meddis R (1988) Simulation of mechanical to neural transduction: further studies. J Acoust Soc Am 83:1056–1063.
Meddis R (2006) Auditory-nerve first-spike latency and auditory absolute threshold: a computer model. J Acoust Soc Am 119:406–417.
Meddis R, O'Mard LPO, Lopez-Poveda EA (2001) A computational algorithm for computing non-linear auditory frequency selectivity. J Acoust Soc Am 109:2852–2861.
Meddis R, Delahaye R, O'Mard LPO, Sumner C, Fantini DA, Winter I, Pressnitzer D (2002) A model of signal processing in the cochlear nucleus: comodulation masking release. Acta Acust/Acustica 88:387–398.
Messing DP, Delhorne L, Bruckert E, Braida LD, Ghitza O (2009) A non-linear efferent-inspired model of the auditory system; matching human confusion in stationary noise. Speech Commun 51:668–683.
Møller AR (1961) Network model of the middle ear. J Acoust Soc Am 33:168–176.
Moore BCJ (2007) Cochlear Hearing Loss. Physiological, Psychological and Technical Issues. Chichester: Wiley.
Moore BCJ, Glasberg BR, Baer T (1997) A model for the prediction of thresholds, loudness and partial loudness. J Audio Eng Soc 45:224–240.
Mountain DC, Hubbard AE (1996) Computational analysis of hair cell and auditory nerve processes. In: Hawkins HL, McMullen TA, Popper AN, Fay RR (eds), Auditory Computation. New York: Springer, pp. 121–156.
Narayan SS, Temchin AN, Recio A, Ruggero MA (1998) Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae. Science 282:1882–1884.
Nedzelnitsky V (1980) Sound pressures in the basal turn of the cat cochlea. J Acoust Soc Am 68:1676–1689.
Nilsson HG (1975) Model of discharge patterns of units in the cochlear nucleus in response to steady state and time-varying sounds. Biol Cybern 20:113–119.
Oono Y, Sujaku Y (1975) A model for automatic gain control observed in the firings of primary auditory neurons. Trans Inst Electron Comm Eng Jpn 58:352–358 (in Japanese) An abstract in English appears in Abstracts of the Trans Inst Elects on Comm Eng Jpn 58:61–62.
Palmer AR, Russell IJ (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair cells. Hear Res 24:1–15.
Pascal J, Bourgeade A, Lagier M, Legros C (1998) Linear and nonlinear model of the human middle ear. J Acoust Soc Am 104:1509–1516.
Patterson RD, Robinson K, Holdsworth J, McKeown D, Zhang C, Allerhand M (1992) Complex sounds and auditory images. In: Cazals Y, Horner K, Demany L (eds), Auditory Physiology and Perception, Oxford: Pergamon, pp. 429–443.
Patterson RD, Unoki M, Irino T (2003) Extending the domain of center frequencies for the compressive gammachirp auditory filter. J Acoust Soc Am 114:1529–1542.
Patuzzi R, Sellick PM (1983) A comparison between basilar membrane and inner hair cell receptor potential input-output functions in the guinea pig cochlea. J Acoust Soc Am 74:1734–1741.
Patuzzi R, Sellick PM, Johnstone BM (1984) The modulation of the sensitivity of the mammalian cochlea by low frequency tones: III. Basilar membrane motion. Hear Res 13:19–27.
Pfeiffer RR (1970) A model for two-tone inhibition of single cochlear-nerve fibers. J Acoust Soc Am 48:1373–1378.
Plack CJ, Oxenham AJ, Drga V (2002) Linear and nonlinear processes in temporal masking. Acta Acust/Acustica 88:348–358.
Plomp R (1976) Aspects of Tone Sensation: A Psychophysical Study. London: Academic.
Recio A, Rich NC, Narayan SS, Ruggero MA (1998) Basilar-membrane responses to clicks at the base of the chinchilla cochlea. J Acoust Soc Am 103:1972–1989.
Rhode WS, Cooper NP (1996) Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo. Audit Neurosci 3:101–121.

Robert A, Eriksson JL (1999) A composite model of the auditory periphery for simulating responses to complex sounds. J Acoust Soc Am 106:1852–1864.
Robles L, Ruggero MA (2001) Mechanics of the mammalian cochlea. Physiol Rev 81:1305–1352.
Robles L, Ruggero MA, Rich NC (1991) Two-tone distortion in the basilar membrane of the cochlea. Nature 349:413–414.
Robles L, Ruggero MA, Rich NC (1997) Two-tone distortion in the basilar membrane of the chinchilla cochlea. J Neurophysiol 77:2385–2399.
Rosowski JJ (1996) Models of external- and middle-ear function. In: Hawkins HL, McMullen TA, Popper AN, Fay RR (eds), Auditory Computation. New York: Springer, pp. 15–61.
Ross S (1982) A model of the hair cell-primary fiber complex. J Acoust Soc Am 71:926–941.
Ruggero MA, Temchin AN (2002) The roles of the external, middle, and inner ears in determining the bandwidth of hearing. Proc Natl Acad Sci U S A 99:13206–13210.
Ruggero MA, Temchin AN (2003) Middle-ear transmission in humans: wide-band, not frequency-tuned? Acoust Res Lett Online 4:53–58.
Ruggero MA, Rich NC, Robles L, Recio A (1990) The effects of acoustic trauma, other cochlear injury, and death on basilar-membrane responses to sound. In: Axelson A, Borchgrevink H, Hellström PA, Henderson D, Hamernik RP, Salvi RJ (eds), Scientific Basis of Noise-Induced Hearing Loss. New York: Thieme, pp. 23–35.
Russell IJ, Murugasu E (1997) Medial efferent inhibition suppresses basilar membrane responses to near characteristic frequency tones of moderate to high intensities. J Acoust Soc Am 102:1734–1738.
Russel IJ, Sellick PM (1978) Intracellular studies of hair cells in the mammalian cochlea. J Physiol 2:261–290.
Sachs MB, Kiang NY (1968) Two-tone inhibition in auditory nerve fibers. J Acoust Soc Am 43:1120–1128.
Schroeder MR, Hall JL (1974) Model for mechanical to neural transduction in the auditory receptor. J Acoust Soc Am 55:1055–1060.
Schwid HA, Geisler CD (1982) Multiple reservoir model of neurotransmitter release by a cochlear inner hair cell. J Acoust Soc Am 72:1435–1440.
Searle CL, Braida LD, Cuddy DR, Davis MF (1975) Binaural pinna disparity: another auditory localization cue. J Acoust Soc Am 57:448–455.
Sellick PM, Russell IJ (1980) The responses of inner hair cells to basilar membrane velocity during low frequency auditory stimulation in the guinea pig cochlea. Hear Res 2:439–445.
Shamma SA, Chadwick RS, Wilbur WJ, Morrish KA, Rinzel J (1986) A biophysical model of cochlear processing: intensity dependence of pure tone responses. J Acoust Soc Am 80:133–145.
Shaw EAG (1966) Earcanal pressure generated by a free sound field. J Acoust Soc Am 39: 465–470.
Shaw EAG (1975) The external ear. In: Keidel WD, Neff WD (eds), Handbook of Sensory Physiology. Berlin: Springer, pp. 455–490.
Siebert WM (1965) Some implications of the stochastic behavior of primary auditory neurons. Kybernetic 2:206–215.
Siegel JH (1992) Spontaneous synaptic potentials from afferent terminals in the guinea pig cochlea. Hear Res 59:85–92
Slaney M (1993) An efficient implementation of the Patterson-Holdsworth auditory filter bank. Apple Computer Technical Report #35. Apple Computer Inc.
Smith RL, Brachman ML (1982) Adaptation in auditory nerve fibers: a revised model. Biol Cybern 44:107–120.
Smith RL, Zwislocki JJ (1975) Short-term adaptation and incremental responses of single auditory-nerve fibers. Biol Cybern 17:169–182.
Smith RL, Brachman ML, Frisina RD (1985) Sensitivity of auditory-nerve fibers to changes in intensity: a dichotomy between decrements and increments. J Acoust Soc Am 78:1310–1316.
Sumner CJ, Lopez-Poveda EA, O'Mard LPO, Meddis R (2002) A revised model of the inner hair cell and auditory nerve complex. J Acoust Soc Am 111:2178–2188.

Sumner CJ, Lopez-Poveda EA, O'Mard LP, Meddis R (2003a) Adaptation in a revised inner-hair cell model. J Acoust Soc Am 113:893–901.
Sumner CJ, O'Mard LPO, Lopez-Poveda EA, Meddis R (2003b) A non-linear filter-bank model of the guinea-pig cochlear nerve. J Acoust Soc Am 113:3264–3274.
Sun Q, Gan RZ, Chang K-H, Dormer KJ (2002) Computer-integrated finite element modeling of human middle ear. Biomechan Model Mechanobiol 1:109–122.
Tan Q, Carney LH (2003) A phenomenological model for the responses of auditory-nerve fibers: II. Nonlinear tuning with a frequency glide. J Acoust Soc Am 114:2007–2020.
von Helmholtz HL (1877) The Sensation of tones. (Translated by AJ Ellis, 1954.) New York: Dover.
Voss SE, Rosowski JJ, Merchant SN, Peake WT (2000) Acoustic responses of the human middle ear. Hear Res 150:43–69.
Walsh T, Demkowicz L, Charles R (2004) Boundary element modelling of the external human auditory system. J Acoust Soc Am 115:1033–1043.
Weiss TF (1966) A model of the peripheral auditory system. Kybernetic 3:153–175.
Westerman LA, Smith RL (1984) Rapid and short term adaptation in auditory nerve responses. Hear Res 15:249–260.
Westerman LA, Smith RL (1988) A diffusion model of the transient response of the cochlear inner hair cell synapse. J Acoust Soc Am 83:2266–2276.
Wiegrebe L, Meddis R (2004) The representation of periodic sounds in simulated sustained chopper units of the ventral cochlear nucleus. J Acoust Soc Am 115:1207–1218.
Wightman FL, Kistler DJ (1989) Headphone simulation of free-field listening: I. Stimulus synthesis. J Acoust Soc Am 85:858–867.
Wilson BS, Schatzer R, Lopez-Poveda EA, Sun X, Lawson DT, Wolford RD (2005) Two new directions in speech processor design for cochlear implants. Ear Hear 26:73S–81S.
Wilson BS, Schatzer R, Lopez-Poveda EA (2006) Possibilities for a closer mimicking of normal auditory functions with cochlear implants. In: Waltzman SB, Roland JT (eds), Cochlear Implants. New York: Thieme, pp. 48–56.
Zeddies DG, Siegel JH (2004) A biophysical model of an inner-hair cell. J Acoust Soc Am 116:426–441.
Zhang X, Carney LH (2005) Analysis of models for the synapse between the inner hair cell and the auditory nerve. J Acoust Soc Am 118:1540–1553.
Zhang X, Heinz MG, Bruce IC, Carney LH (2001) A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. J Acoust Soc Am 109:648–670.
Zwislocki J (1962) Analysis of the middle-ear function. Part I: Input impedance. J Acoust Soc Am 34:1514–1523.

http://www.springer.com/978-1-4419-1370-8

Computational Models of the Auditory System
Meddis, R.; Lopez-Poveda, E.; Fay, R.R.; Popper, A.N. (Eds.)
2010, XII, 350 p., Hardcover
ISBN: 978-1-4419-1370-8