HEARING LOSS AND FREQUENCY ANALYSIS OF COMPLEX SOUNDS

Marjorie Leek and Michelle Molis
National Center for Rehabilitative Auditory Research
Portland VA Medical Center
Portland, Oregon 97239

The sounds that surround us in our everyday lives range from very simple and tonal, such as one might hear from a flute or a whistle, to highly complex containing multiple tones and noises, such as listening to a single speaker in the presence of background noise or the babble of many other talkers at the same time. For the most part, the sounds that are relevant to us are complex, which, by definition, means they are made up of multiple frequencies with multiple sound levels. Some may be tonal, some may have a noisy quality, and often there will be important temporal features such as sequence or order effects. Speech is just such a signal—perhaps the most relevant and complex signal heard by humans. The auditory system is exquisitely designed to encode information from these complex sounds. Encoding, combining, and recombining information allows us to sort out our sound environment and gain information about who (or what) is near or in the distance, who is speaking, if we are in danger, or if there is something good to eat out there. The study of psychoacoustics and physiology allows us to understand what those encoding mechanisms are, how they work separately and together, and how they might fail us from time to time.

Although the sense of hearing is traditionally associated with the ears, hearing is actually accomplished in the brain, after acoustic information has been transmitted through the peripheral auditory system and coded throughout the auditory nervous system to combine information from the two ears across frequency and time. For this processing to be complete and accurate, leading to successful perception of the intended acoustic signal, auditory structure and physiological function must be intact, or nearly so.

Hearing loss is a significant barrier to everyday communication and represents a financial, psychological, economic, and vocational burden on millions of people around the world. According to the National Institute on Deafness and Other Communication Disorders (NIDCD), some 36 million adults in the United States alone experience hearing loss, and that number is growing as the population ages (http://www.nidcd.nih.gov). It is rather rare for individuals to have no hearing at all, and much more common to have reduced hearing sensitivity ranging from mild to more severe hearing loss. It is not unusual for an individual to have some frequency regions of normal or near-normal hearing accompanied by some regions where hearing is severely impaired.

"There is considerable research showing that spectral analysis is impaired in the presence of sensorineural hearing loss, and this loss of frequency resolution is influential in the difficulty understanding speech in noisy environments."

Each of these forms of partial or, in some cases, nearly complete loss of hearing results in changes in the encoding mechanisms that feed acoustic information to the brain. Some of these coding deficits can be partially overcome with treatments, most commonly in the form of either a hearing aid or a cochlear implant. But these treatments can bring with them their own characteristic distortions and signal degradations that combine with impaired auditory function to make everyday listening difficult.

The most common source of hearing loss in adults occurs as the result of damage to the hair cells in the inner ear, which provide the initial stages of processing (outer hair cells) as well as stimulation to the auditory nerve (inner hair cells). The hair cells are arrayed longitudinally in four rows on the basilar membrane stretching along the length of the cochlea. In humans, the length of the basilar membrane is about 30 mm, and it runs from the cochlear base, at the interface between the middle ear and the inner ear, to the apex, which is at the tip of the cochlear spiral. A sound wave that is transmitted to the inner ear will be distributed in frequency along the length of the basilar membrane, with different frequencies providing stimulation in different regions along the length of the cochlea. The location of maximum stimulation along the basilar membrane is tonotopically organized according to frequency, with higher frequencies encoded near the base of the cochlea and lower frequencies stimulating maximally at the apex of the cochlea. The tonotopic arrangement is maintained throughout the auditory pathways up to and including the primary auditory cortex.

While the inner hair cells provide stimulation to the auditory nerve, the three rows of outer hair cells do not transmit auditory information, but instead function as sound processors to enhance auditory perception. These hair cells respond in a nonlinear manner to different levels of stimulating sound, amplifying low to moderate levels, but not providing gain at high levels. The nonlinearity of level processing accounts for the wide dynamic range of sound levels that we can hear, and it also supplies the sharp frequency tuning of the auditory response in normally functioning ears.

The outer hair cells are particularly vulnerable to damage from excessive sound levels (i.e., noise exposure) and other sources of hearing loss such as disease states, ototoxic drugs, and the aging process. Damage or destruction of outer hair cells disrupts nonlinear processing in the cochlea, resulting
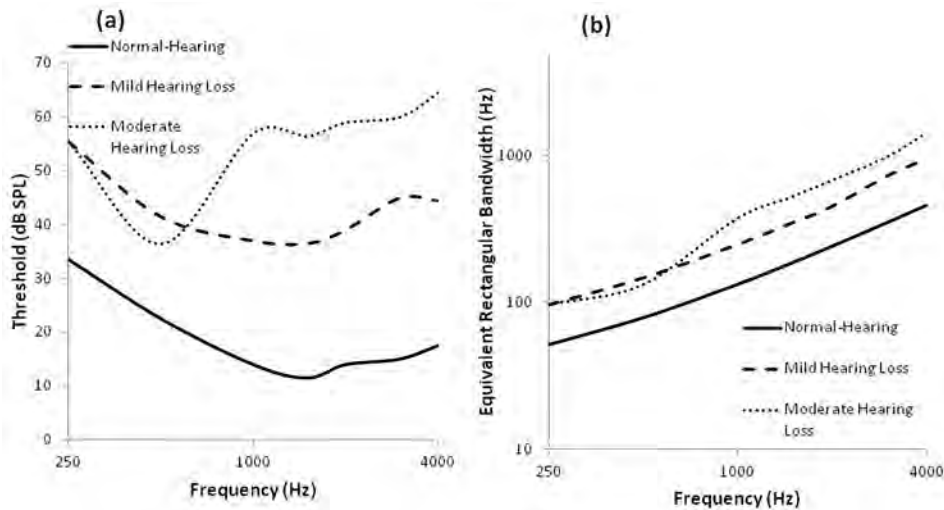
Fig. 1. (a) Auditory thresholds in quiet for normal hearing-listeners (solid line), a listener with a mild hearing loss (dashed line), and a listener with a moderate hearing loss (dotted line). (b) Auditory filter bandwidths estimated for the normal-hearing, and two listeners with hearing loss at several frequencies.

in a loss of auditory sensitivity, a reduction in the dynamic range of sound levels we hear, a loss of normal loudness processing, and impairments to the normal frequency selectivity, which serves to separate sounds into bands of frequency components. Each of these functional effects of the loss of outer hair cells has a significant impact on the manner in which sounds are encoded for perception by the brain.

In people with hearing impairment, the mechanisms that supply the nonlinear amplification have been lost or damaged, and therefore, although their perception of very loud sounds might be relatively normal, they do not enjoy the internal gain provided for soft sounds. This is the underlying cause of the increased audiological thresholds which define a hearing loss clinically. However, it should be remembered that for many hearing-impaired people, complex, wideband sounds (such as speech) typically span regions of normal hearing and frequency regions of impaired hearing, so that the combined stimulation to the brain is not only different from normal processing locally, but the composite that occurs through processing channels is distorted compared to the combined outputs of frequency channels in wholly normal auditory systems. Figure 1a shows a frequency threshold curve for average normal hearing listeners, along with two additional threshold curves, one depicting a rather mild hearing loss (dashed line), and one showing a more severe loss (dotted line). Notice that although in both impaired cases, the thresholds in the lower frequencies are not much different from the normal curve, as is fairly common, hearing loss becomes more severe in the mid to upper frequencies.

In the initial stage of cochlear processing of complex sounds a frequency analysis separates the incoming broadband signal into overlapping bands or channels. Most normal-hearing listeners have channels that are about 10% of the center frequency in bandwidth, i.e., the absolute bandwidths increase with increasing frequency. The nonlinear gain provided by the normal cochlea makes these frequency channels fairly narrowly tuned at low amplitude levels, but the channels become broader and more overlapping at high sound amplitudes. These auditory channels or bandwidths are broadened in the presence of cochlear damage. The frequency channels defined by the auditory filters in the impaired cochlea are broader and more overlapping than is observed in normal-hearing listeners because of the loss of the nonlinearity in cochlear processing. Although there is not a perfect correspondence, broader bandwidths may be somewhat predicted from the degree of hearing loss.

Figure 1b shows an estimate of the bandwidths (in Equivalent Rectangular Bandwidths–ERB) for normal-hearing listeners at a fairly low level (solid line). Glasberg and Moore (1990) developed an equation that fits the measured auditory filter bandwidths for these listeners fairly closely. The other two lines in this panel (dashed and dotted) are estimates of the auditory filter bandwidths calculated from the thresholds for the two hearing-impaired subjects in Fig 1a, based on calculations described in Moore and Glasberg (2004).

One way of visualizing the functioning of the normal and abnormal auditory filter banks is by constructing an excitation pattern. This analysis combines the hypothesized cochlear outputs across frequency for a given stimulus, and is meant to simulate the spectral representation after processing through a bank of auditory filters. There are multiple potential effects of the impaired frequency resolution that occurs in hearing-impaired listeners as a result of damage to outer hair cells. One of these is a smearing of the peaks and valleys that provide information about the spectrum of complex sounds like speech. Figure 2 shows the amplitude spectrum of the vowel /a/, as in "ball." Note the characteristic peaks in this spectrum, the formants, separated by valleys. The frequency locations of the peaks in each vowel sound are the primary sources for identification of these sounds in speech. When a complex spectrum such as this is processed by a bank of auditory filters, some of the detail in the separation of peaks and valleys gets reduced. However, in normal-hearing listeners, sufficient peak-to-valley differences are maintained such that the processed vowel sound can lead to accurate identification.

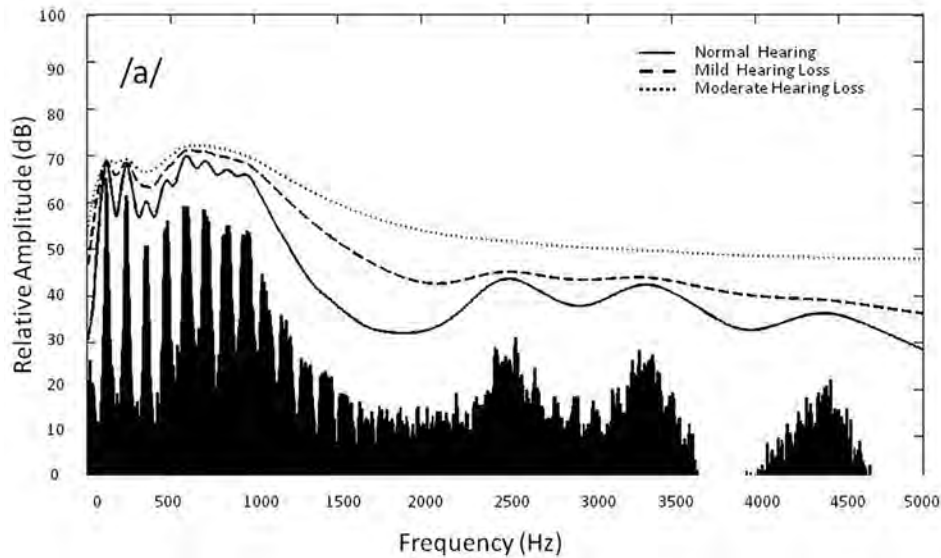Figure 2 demonstrates how the processing through the

Fig. 2. Spectrum for the vowel /a/ and excitation patterns calculated for normal-hearing listeners and two listeners with hearing loss. Excitation patterns indicate good preservation of formant peaks in the normal system, but much impoverished representation of higher formants in the impaired systems.

bank of filters may differ between a normal cochlea and two different simulated severities of cochlear damage. The solid black line represents the normal excitation pattern in response to the vowel sound shown in the figure. The peaks at low frequencies show the response to the individual harmonics of the vowel sound because the filter bandwidths at lower frequencies are narrow, and only one or two harmonics of the vowel are processed through each filter. The overall peaks around 600 Hz, 2600 Hz, and 3300 Hz represent the first three formant or resonant frequencies that identify the sound. The dashed line displays an excitation pattern generated by simulating a filter bank with twice normal bandwidth filters, and the dotted line shows three times normal filter bandwidths. Typically, for a moderate-to-severe hearing loss, auditory filters may be from two to four times broader than in a normal cochlea (Moore and Glasberg, 2004).

The excitation pattern for the normal-hearing system shows clear peaks and valleys that represent the formants that characterize this particular vowel sound. To the extent that these peaks and valleys are less precisely represented in the auditory system, the speech sounds will be less well identified

and there will be more ambiguity in the speech percept. This is demonstrated in the top two excitation patterns, showing representations of the vowel in the case of mild and moderate hearing losses. With each broadening in the filter bandwidths, the excitation pattern becomes flatter, with less definition of the peaks and valleys in the vowel spectrum. As the peaks are less well-defined the vowels are perceived with much greater ambiguity because a hearing-impaired listener may not accurately process the spectral detail (Molis and Leek, 2011).

In addition to a smearing of the peaks and valleys in the internal excitation pattern, the broader auditory filters in hearing-impaired listeners interfere with normal perception in yet another way. The excitation pattern calculation only uses the auditory spectrum, but the input to the ear is always a temporal waveform, not an amplitude spectrum. The consequences of separating a complex waveform into frequency channels, but preserving the temporal characteristics of the channels, rather than limiting the output to only the spectrum, may be seen using a computer model of auditory processing developed by Patterson and his colleagues (1995)

/a/

Formant 3 → 3342 Hz
3069 Hz
Formant 2 → 2584 Hz
992 Hz
Formant 1 → 625 Hz
205 Hz
113 Hz

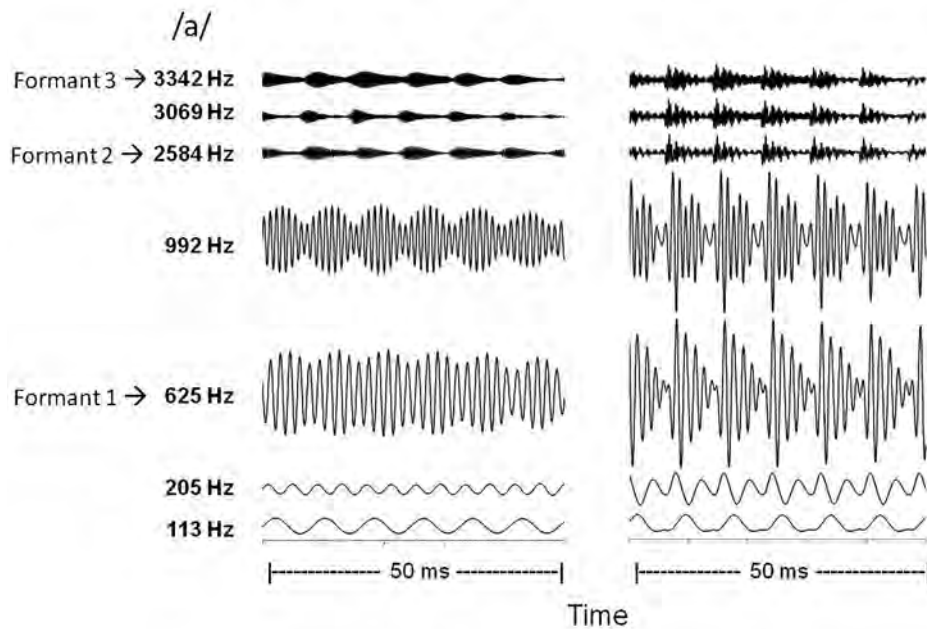|← ---------- 50 ms ---------- →| |← ---------- 50 ms ---------- →|

Time

Fig. 3. Outputs of a bank of normal auditory filters (left panel) and a bank of filters with three times the normal width (right panel) for the spoken vowel /a/. The channels centered at the same frequencies are shown for both filter banks. These channels were selected to show the first three formant frequencies, as well as some non-formant channels. Note the greater complexity of the simulated hearing-impaired outputs for all channels.

called the Auditory Image Model (AIM). The model developed for normal-hearing individuals may be modified to change the normal auditory channel bandwidths and view the hypothesized temporal outputs of selected channels simulating both a normal filter bank and one with abnormally broad filters. Figure 3 shows seven channels of output for the vowel /a/ (seen in spectral form in Fig. 2) for a normal-width bank of auditory filters on the left, and for filters that are three times the normal bandwidth on the right, simulating a moderate-to-severe hearing loss. The output of each auditory filter provided to its associated auditory nerve fibers may be visualized as a waveform, constructed of combinations of frequencies that are passed by each filter. The channels shown were selected to be those with center frequencies at the first, second, and third formants of the vowel /a/, as well as four intermediate non-formant frequency channels. The narrow filter bandwidths, found at lower frequencies in a normal cochlea, will pass only a few separate frequencies, and thus the output of those narrow filters will be rather simple, resembling pure tones. At higher frequency regions, more component frequencies are passed by the filters that become progressively broader, and therefore the outputs of those channels are more complex as more frequencies interact.

These channel outputs which combine component frequencies in the stimulus, provide complementary information to the auditory system about the input sound. The hypothesized channel outputs in a hearing-impaired cochlea, with broader auditory filters, become distorted in characteristic ways: first, there are fewer channels that output simple waveforms because more frequency components are passed through impaired channels relative to normal channels for the same frequency location, and second, the production of the more highly complex output waveforms occurs at lower-

than-normal frequency regions. In addition, the broader filters in the impaired ear have greater overlap with each other, further distorting the processed sound. The increased complexity for the hearing-impaired model outputs indicates more interactions among frequency components, perhaps leading to greater distortion as this impaired version of cochlear outputs is encoded for passage into the brain. This demonstrates not only a spectral distortion as seen in the excitation pattern, but also a temporal waveform distortion at the outputs of the frequency channels, which likely also interferes with the accurate recognition of speech.

The analysis of complex sounds provided in the cochlea is obviously not limited only to the frequency analysis described here. However, there is considerable research showing that spectral analysis is impaired in the presence of sensorineural hearing loss, and this loss of frequency resolution is influential in the difficulty understanding speech in noisy environments (ter Keurs *et al.*, 1992). Additional impairments to auditory encoding, such as temporal processing and impaired loudness perception, are likely also involved in the distortions experienced by listeners with hearing loss. In fact, there is growing evidence that the reduction of fine temporal precision with hearing loss may be a significant problem that underlies both the extraction and understanding of speech in noise, as well as the poor pitch perception manifested in an inability of hearing-impaired people to enjoy music (Lorenzi *et al.*, 2006; Moore, 2008).

The effects of impaired auditory processing result in difficulties extracting a target speech signal from among many other talkers. Hearing loss may also make it difficult to recognize voices and to localize sounds in the environment. All of these impaired functions correspond to the auditory experiences described by people with hearing loss, who complain

that not only are sounds too soft, but that they are indistinguishable and in some cases undecipherable.**AT**

References

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hearing Res. **47**, 103–138.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. (**2006**). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," Proc. Natl. Acad. Sci. **103**, 18866–18869.

Molis, M. R., and Leek, M. R. (**2011**). "Vowel identification by hearing-impaired listeners in response to variation in formant frequencies," J. Speech-Language-Hearing Res. **54**, 1211–1223.

Moore, B. C. (**2008**). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," J. Assn. Res. Otolaryngol. **9**, 399–406.

Moore, B. C. J., and Glasberg, B. R. (**2004**). "A revised model of loudness perception applied to cochlear hearing loss," Hearing Res. **188**, 70–88.

Patterson, R. D., Allerhand, M. H., and Giguère, C. (**1995**). "Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform," J. Acoust. Soc. Am. **98**, 1890–1894.

ter Keurs, M., Festen, J. M., and Plomp, R. (**1992**). "Effect of spectral envelope smearing on speech reception. I," J. Acoust. Soc. Am. **91**, 2872–2880.

Marjorie Leek is a Senior Research Career Scientist at the National Center for Rehabilitative Research, Portland VA Medical Center, and Professor in the Department of Otolaryngology Head/Neck Surgery at the Oregon Health and Science University. She received her Ph.D. in 1980 in Hearing Science from the University of Kansas. Her research focuses on auditory perception of complex sounds and speech by normal-hearing and hearing-impaired individuals.

Michelle Molis received her Ph.D. in 2002 in experimental psychology from the University of Texas at Austin. She joined the National Center for Rehabilitative Auditory Research at the Portland VA Medical Center in 2005 as an Investigator. Her research focuses on speech perception by people with hearing loss as well as normal-hearing individuals, and in modeling vowel perception. She is particularly interested in understanding the distortions in vowel sounds that result from processing by an impaired auditory system.