Department of Veterans Affairs

Journal of Rehabilitation Research and Development Vol. 38 No. 1, January/February 2001
Pages 111–121

Noise reduction in hearing aids: a review

Harry Levitt, PhD
The Lexington School for the Deaf/Center for the Deaf, 75th Street and 30th Avenue, Jackson Heights, NY 11370

Abstract—Background noise is particularly damaging to speech intelligibility for people with hearing loss. The problem of reducing noise in hearing aids is one of great importance—and great difficulty. The problem has been addressed in many different ways over the years. The techniques used range from relatively simple forms of filtering to advanced signal processing methods. This paper provides a brief overview, in nontechnical language, of the issues involved and the various approaches to solving the problem.

Key words: *adaptive noise cancellation, digital signal processing, directional microphones, hearing aids, noise reduction, speeach-in-noise.*

INTRODUCTION AND GENERAL PRINCIPLES

It is well known that background noise reduces the intelligibility of speech and that the greater the level of background noise the greater the reduction in intelligibility. We are able to understand speech in a moderately noisy environment because speech is a highly redundant signal and thus even if part of the speech signal is masked by noise, other parts of the speech signal will convey sufficient information to make the speech intelligible, or at least sufficiently intelligible to allow for effective speech communication. There is less redundancy in the speech signal for a person with hearing loss since part of the speech is either not audible or is severely distorted because of the hearing loss. Background noise that masks even a small portion of the remaining, impoverished speech signal will degrade intelligibility significantly because there is less redundancy available to compensate for the masking effects of the noise. As a consequence, people with hearing loss have much greater difficulty than normally hearing people in understanding speech in noise (1–3).

Hearing aids allow for some degree of signal processing to reduce the effects of noise. The recent development of digital hearing aids opens up substantial new possibilities with respect to the use of advanced signal-processing techniques for noise reduction (4,5). Because of the particularly damaging effects of background noise on speech intelligibility for people with hearing loss (i.e., hearing-aid users) this problem is of critical importance.

The general problem of noise reduction is not new and has been addressed in great depth by statisticians, physicists, engineers, and others (6–8). The problem is central to

Preparation of this paper was supported by Grant H133E980010, Rehabilitation Engineering Research Center on Hearing Enhancement, U.S. Department of Education/OSERS/NIDRR, National Institute on Disability and Rehabilitation Research.

Address all correspondence and requests for reprints to: Harry Levitt, PhD, 998 Sea Eagle Loop, PO Box 610, Bodega Bay, CA 94923; email hlevitt@pon.net

the fields of Information Theory and Coding Theory. As a consequence, there is a substantial body of theory and methods of practical implementation that address the problem. Unfortunately, the problem is fundamentally very difficult for the most common types of noise and there are severe limits as to how much noise reduction is in fact possible. On the positive side, there are special considerations relating to the joint effects of hearing loss and background noise on speech intelligibility that allow for the development of signal-processing strategies that may be of benefit to the hearing-aid user. The objective in developing these techniques is not so much to reduce background noise, *per se*, but to reduce the effects of background noise on speech intelligibility and overall sound quality (9,10). The purpose of this paper is to review the issues involved in order to provide a realistic picture of what can be done and likely future developments.

It is first necessary to define what is meant by noise. Noise is any unwanted signal that interferes with a desired signal. Speech is the signal of primary interest in this discussion and there are three types of noise that are particularly damaging to speech intelligibility:

1. Random noise with an intensity-frequency spectrum similar to that of speech.
2. A second interfering voice. Note that the interference produced by many other voices of roughly equal intensity (known as speech babble) has physical characteristics similar to that of random noise with a speech-shaped intensity-frequency spectrum.
3. Substantial room reverberation. Reverberation is produced by sound being reflected off walls, floors, ceilings, table tops, and other hard surfaces. Some reverberation is helpful in reinforcing the speech signal, but too much reverberation will reduce speech intelligibility (and overall sound quality), particularly in the presence of other types of noise.

General principals are very helpful both for specifying the nature of a problem and for identifying possible ways of addressing the problem. The following two general principles apply to the problem of speech and noise: The more we know about the speech and noise, the more we can do to reduce the effects of the noise on the speech; and, the larger the differences between the speech and the noise, the more we can do to reduce the effects of the noise on the speech.

In order to apply these principles to the problem of noise reduction in hearing aids, it is necessary to have a basic understanding of how the auditory system processes sound, and the effect of a hearing impairment on this processing. An in-depth review of auditory signal processing is not possible in the space available, but two salient issues need to be mentioned.

First, the peripheral auditory system analyzes sound by means of a bank of overlapping, narrowband filters. These filters are known as the critical bands of hearing. The exact shape and width of these filters is still the subject of much research (11–14). For the purposes of this discussion, however, we can think of these filters as being similar to a bank of 1/3 octave-band filters, but of slightly narrower bandwidth. Unlike a bank of contiguous filters, however, the critical bands are asymmetric, with substantial overlap. As a consequence, a critical band centered on a higher frequency will pick up a low-frequency sound. Noise in a critical band will thus not only mask signals in that critical band but will also mask, to a lesser extent, signals in higher-frequency bands (15–17). This effect, known as upward spread of masking, is relatively small at low noise levels but increases with increasing noise level. It can be quite large at very high levels, such as that resulting from high-gain amplification of relatively intense background noise.

Secondly, most hearing-aid users have sensorineural hearing loss. A major problem in providing amplification for this type of hearing loss is that the dynamic range of hearing is reduced. Not only is the threshold of hearing raised as a result of the hearing loss, but the threshold of loudness discomfort remains the same or is often even lower than that of normal hearing. Most sensorineural hearing losses show an elevation of the hearing threshold that increases with increasing frequency. As a consequence, the dynamic range of hearing (from threshold to discomfort) is usually much narrower in the high frequencies (18,19).

Any practical method of noise reduction must take the above factors into account. One way of dealing with the reduced dynamic range of hearing is to limit the output of the hearing aid in some way. This is necessary because the gain that is needed to make the weaker sounds of speech audible will at the same time make the stronger sounds uncomfortably loud. A simple way of limiting the output of a hearing aid is to clip the peaks of the amplified signal whenever they exceed a critical level (e.g., the loudness discomfort level). Peak clipping, however, introduces substantial nonlinear distortion. An approach that produces relatively little distortion is to automatically reduce the gain of the hearing aid substan-

tially when the amplified sound reaches a critical output level. This approach is known as compression limiting and is being used increasingly in modern hearing aids. A third approach to limiting the output of a hearing aid is to reduce the gain progressively as the level of the input signal increases. This technique, known as wide dynamic range (WDR) amplitude compression, can be used to match the characteristics of the hearing loss not only with respect to threshold shift and maximum acceptable loudness, but also with respect to the increase of loudness with intensity within the dynamic range of hearing.

A very useful tool for analyzing the effects of masking, filtering, and hearing loss on speech intelligibility is the Speech Interference Index (SII; reference 20). This index is essentially a weighted average of the speech-to-noise ratios in a set of frequency bands that approximate the critical bands of hearing. The level of the speech peaks relative to the rms level of the noise (in dB) is used to determine the speech-to-noise ratio in each frequency band. Negative speech-to-noise ratios are assigned a weight of zero since there is no contribution to intelligibility if the noise exceeds the speech in any frequency band. There is also a maximum contribution to intelligibility for each frequency band. The SII includes the effect of hearing loss by taking the threshold of hearing into account. For example, if the speech level is below the threshold of hearing in any frequency band, then the contribution to intelligibility is zero for that band.

The sections that follow provide illustrative examples of how the general principles cited above are applied to the problem of noise reduction in hearing aids. The first example deals with the relatively simple case of a time-invariant noise that differs in spectral shape from that of speech. Subsequent examples deal with more difficult problems, such as that of time-varying noises as well as the complex spectro-temporal characteristics of speech.

FIXED FILTERS AND TIME-INVARIANT NOISE

Figure 1 shows a typical intensity-frequency spectrum of speech (averaged over time) and a typical intensity-frequency spectrum of a steady-state ambient noise. The speech and noise spectra differ substantially and it is possible to eliminate much of the noise and a relatively small portion of the speech by means of a high-pass filter. Assume, for the purpose of this discussion, that the high-pass filter attenuates all signals below 1 kHz and passes without attenuation all signals above 1 kHz. According to the SII, all frequency bands below 0.4 kHz have a negative speech-to-noise ratio and make no contribution to intelligibility. Eliminating both speech and noise in this frequency region will have no effect on intelligibility since the speech is already masked by the noise. The noise components in this frequency region are also the most intense, and eliminating these components has the desired effect of reducing the loudness of the noise and improving overall sound quality.

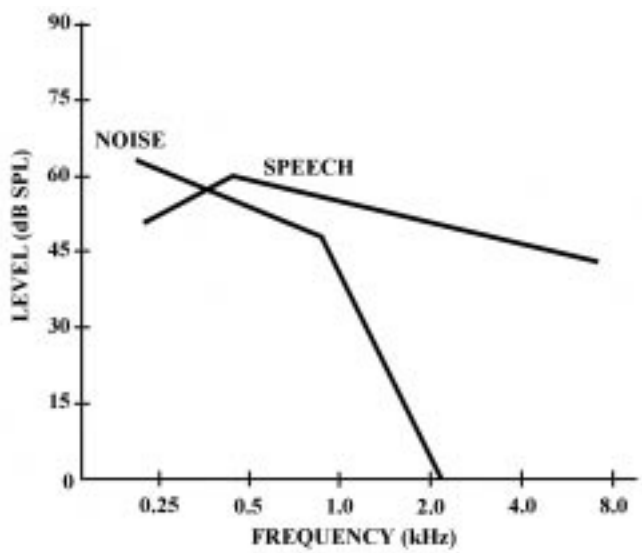

Figure 1.
Long-term spectra of speech peaks and steady-state noise.

The high-pass filter, however, also eliminates both speech and noise in the frequency region between 0.4 kHz and 1.0 kHz. In this region, the frequency spectrum of the speech is slightly above that of the noise, so a small contribution to the SII (and hence to the intelligibility of the speech) is lost. At the same time, the loudness of the noise is reduced further so that there is a tradeoff between improved overall sound quality (e.g., a less-annoying noise) for a small reduction in intelligibility.

In summary, the differences between the speech and the noise lie primarily in the shape of their frequency spectra. Since most of the noise power is concentrated in the low frequencies, the speech is masked in this frequency region and filtering out both speech and noise over this frequency range will have little or no effect on intelligibility but will reduce the loudness and annoyance of the noise; i.e., overall sound quality will be improved.

If, however, the high-pass filter eliminates frequency regions in which the speech-to-noise ratio is positive, even by a small amount, there will be some loss of intelligibility. It is thus of critical importance to match the frequency response of the filter to the spectral characteristics of the noise.

ADAPTIVE FILTERS AND TIME-VARYING SPECTRA

The example of the preceding section is highly idealized. The frequency spectra of everyday noises are seldom so different from that of speech and are sufficiently time invariant that a fixed high-pass filter can effectively eliminate most of the noise without reducing speech intelligibility at the same time (21). It is possible to use adaptive filtering (or frequency-dependent amplitude compression) to reduce noise levels without a significant reduction in intelligibility. The method is to obtain an estimate of the noise spectrum in some way and then to attenuate those frequency bands in which the noise exceeds the speech (22,23). This approach can also be used to reduce reverberation by identifying the frequency bands with excessive reverberation and then attenuating those bands (24).

A practical problem in implementing the above approach is that of obtaining a reasonably accurate estimate of the noise spectrum as it varies over time. One approach to this difficult problem is to measure the noise spectrum during pauses or other short breaks in the speech signal. Since this noise spectrum is obtained over a short interval of time, it is known as the short-term noise spectrum. It is assumed that the short-term noise spectrum does not vary rapidly with time and an appropriate frequency-gain characteristic is then chosen for the speech plus noise, when the speech is once again present.

The mathematical theory of filtering provides a formula for an optimum filter that will maximize the signal-to-noise ratio (25). This filter, known as a Wiener filter, requires that the spectra of both the signal and the noise do not vary with time—a requirement that clearly does not apply to speech. Many speech sounds, however, have spectra that are approximately constant over short intervals of time. It is thus possible to use a short-term Wiener filter in which the short-term spectra of the speech and the noise are assumed not to vary significantly over short intervals of time. The potential gain in the speech-to-noise ratio, assuming the validity of this assumption, is relatively small. The gain in signal-to-noise is obtained separately for each critical band and is small because of the relatively narrow bandwidths. Short-term Wiener filtering for speech in random noise has thus far not proven successful for people with normal hearing. There is some evidence, however, that some people with sensorineural hearing loss and relatively large critical bands may benefit from short-term Wiener filtering, provided the short-term speech and noise spectra are obtained reliably (26). Research in this area is still active.

A variation of the above approach is to take the short-term noise spectrum obtained during a pause in the speech and subtract it from the speech-plus-noise spectrum when speech is once again present (27,28). This technique takes into account time-varying changes in the short-term speech spectrum but still assumes that the short-term noise spectrum does not vary significantly with time. The technique, known as spectrum subtraction, can improve speech-to-noise ratios for many commonly encountered ambient noises by as much as 10 or 12 dB, but without a concomitant improvement in speech intelligibility. This is because the signal processing involved produces audible distortions, referred to as processing noise, that counteract the potential improvements in intelligibility resulting from the reduction in background noise. For listeners who prefer low-level processing noise to high-level random noise, the technique provides an improvement in sound quality with no significant change in intelligibility.

Another approach to the problem, which produces a much-improved speech-to-noise ratio and improved sound quality but no significant change in intelligibility, is that of sinewave modeling (29,30). In this case, the major peaks in the speech-plus-noise spectrum are obtained. These peaks, which are frequently located at the harmonics of voiced speech sounds, consist mostly of speech, with relatively little noise. The spectral components between these peaks, which consist mostly of noise, are discarded. The spectral peaks are then converted back to a time waveform with a much-improved speech-to-noise ratio (approaching 12 dB for speech in white noise), but with some processing noise.

There have been many attempts over the years at reducing background noise and improving the speech-to-noise ratio based on spectral differences between the speech and the noise. Most of these techniques have been variations of the techniques described above, the best of which have yielded essentially the same result: improved speech-to-noise ratio, some processing noise, and no sig-

nificant change in speech intelligibility (9,10). There have also been attempts at improving perceived differences in the speech-to-noise ratio by focusing on differences in the temporal properties of speech and noise. In this case, various combinations of amplitude compression and expansion with carefully chosen time constants effectively reduce noise levels during pauses in the speech, or when the speech signal is relatively weak.

REDUCING SPREAD OF MASKING

Advanced signal-processing techniques have had relatively little success in improving speech intelligibility in random noise for people with normal hearing. There are, however, some conditions under which speech intelligibility in noise can be improved for people with hearing loss. Most of the positive results that have been obtained, thus far, have been under carefully controlled laboratory conditions, but they do point the way to the development of improved hearing aids for noise reduction. The main focus of these techniques is to reduce spread-of-masking effects (16,17,31). These effects are typically greater for hearing-aid users because of the high sound levels resulting from high-gain amplification.

Significant spread of masking will occur under the following rather special conditions.

1. The interfering noise must be relatively intense over a limited frequency region with a rapidly falling spectrum at the boundaries of the noise band.
2. The filter that is used must attenuate only those frequency bands in which the noise exceeds the speech.
3. If the noise spectrum varies with time, an adaptive filter is required that can accurately track the variations in the short-term noise spectrum.

Figure 2 shows a narrow band of noise of high intensity in the low frequencies. Most of the noise is concentrated in the frequency region below 0.25 kHz, the frequency spectrum falling off sharply with increasing frequency above 0.25 kHz. This intense low frequency noise not only masks weaker signals in the region below 0.25 kHz but will also mask signals at higher frequencies, as shown schematically by the dashed line in **Figure 2**. This dashed line represents the upward spread of masking produced by the high intensity, low frequency noise. The SII for the speech and noise spectra shown in the figure can be derived by simply treating the dashed line as the effective masking spectrum of the noise. Spread of masking increases with increasing noise level and thus by attenuating the noise, either by filtering or amplitude compression, spread of masking is reduced, resulting in an improved SII which, in turn, should result in an increase in intelligibility.

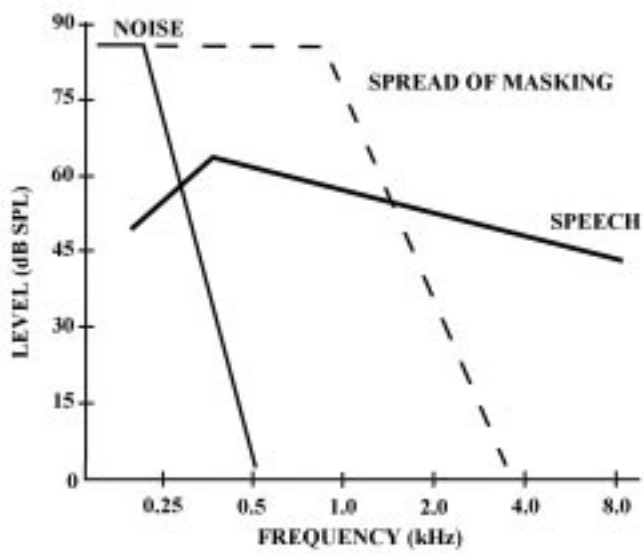

Figure 2.
Upward spread of masking produced by an intense low-frequency band of noise.

Unfortunately, for the types of noise encountered in everyday life, the predicted increase in intelligibility is relatively small and is often offset by limitations in implementing the appropriate method of signal processing. It is important to bear in mind that for a hearing aid to be practical it must be cosmetically acceptable; i.e., it must be extremely small and not noticeable. The most popular hearing aids today are small enough to fit inside the ear canal and are barely visible to the naked eye. The development of signal-processing hearing aids of such small size is a remarkable engineering achievement, but there is a price to be paid. Amplification systems of extremely small size and powered by a low voltage source (e.g., a hearing-aid battery) are subject to relatively high levels of internal noise and nonlinear distortion. In addition, the close proximity of the input microphone to the output transducer (loudspeaker) results in unstable acoustic feedback under high gain conditions, thereby adding another significant constraint to an already difficult engineering problem.

There have been several attempts at developing signal-processing hearing aids for improving speech intelli-

gibility by reducing spread-of-masking effects. The earliest hearing aids designed to improve speech intelligibility by reducing spread of masking were not successful, however, largely because the potential gains in intelligibility are small and because imperfections in the electroacoustic characteristics of these instruments (as a result of the constraints imposed by small size, low voltage, low power consumption) have a greater effect in reducing speech intelligibility than the gains that could, in principle, be realized from a reduction in spread of masking (32–34). Recent advances in the micro-miniaturization of digital signal-processing chips have allowed for the development of a new generation of signal-processing hearing aids with improved electroacoustic characteristics and much greater flexibility in implementing adaptive frequency-gain characteristics. Significant improvements in overall sound quality can be obtained with these instruments if fitted properly, with possibly a small improvement in speech intelligibility under specific conditions; e.g., intense background noises with power concentrated in narrow frequency bands.

SPATIAL FILTERING

Speech and noise can differ not only in their spectral and temporal properties, but also in their spatial properties. It is possible to make good use of spatial differences to improve speech intelligibility using directional microphones or microphone arrays. There are, however, important limitations on how much separation can be achieved in practice, as is evident from the following example.

Figure 3a shows speech and noise reaching a hearing-aid microphone from two different directions. Both the speech and the noise are generated in an anechoic room so that there are no reflections. If an omnidirectional microphone is used (i.e., a microphone that picks up sound from all directions), both speech and noise will be picked up simultaneously and there will be a noise interference problem.

If a directional microphone is used (i.e., a microphone that picks up sound from one direction only) it is possible, in principle, to pick up the speech and eliminate the noise, as shown in **Figure 3b**. The heavy lines in the diagram identify the range of directions from which the microphone will pick up sound. This is, of course, a very useful approach but it does not always work well because microphones that are small enough to fit on a hearing aid have limited directional capabilities, particularly in the low frequencies, and cannot separate speech and noise as effectively as shown in **Figure 3b**. There is also the problem of room reverberation.

We typically listen to speech and noise in a room, and rooms typically have walls, and walls reflect sound.

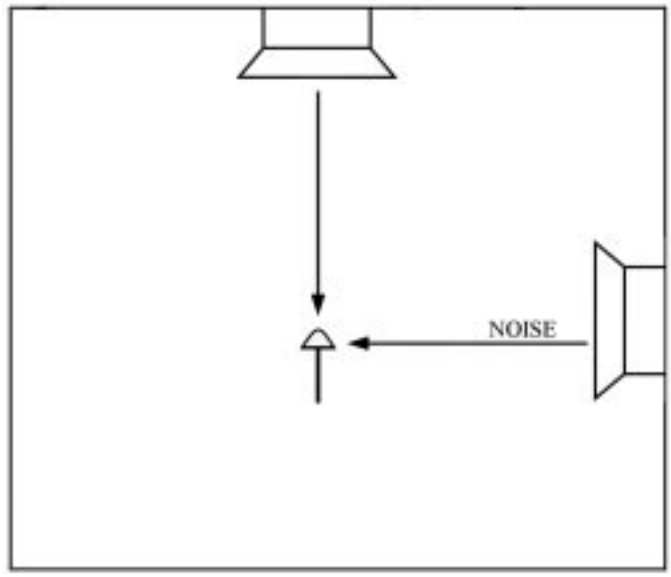

Figure 3a.
Speech and noise reaching a microphone from two different directions.

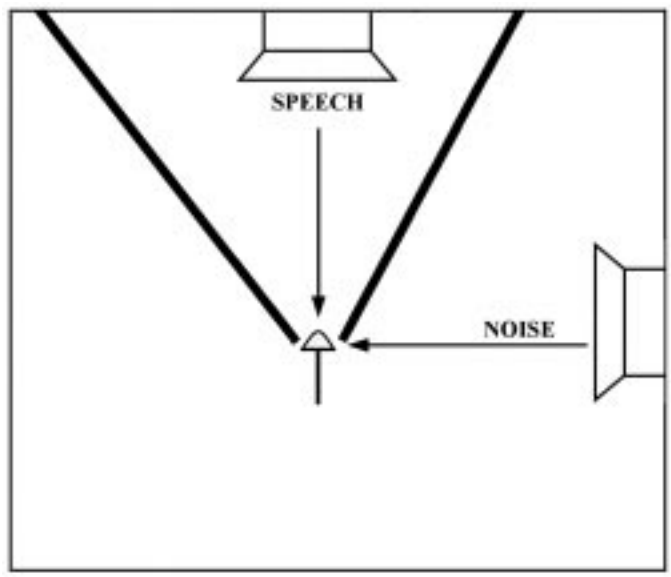

Figure 3b.
Attenuation of background noise by a directional microphone under idealized, non-reverberant conditions. The heavy lines show the range of directions within which sound is picked up by the microphone without attenuation.

Figure 3c shows an example of a single reflection. Some of the noise is reflected off the top wall, reaching the microphone after being reflected only once. Sound can thus reach the microphone in two different ways, by direct transmission from source to microphone and by reflections off walls and other surfaces. This is true for both speech and noise.

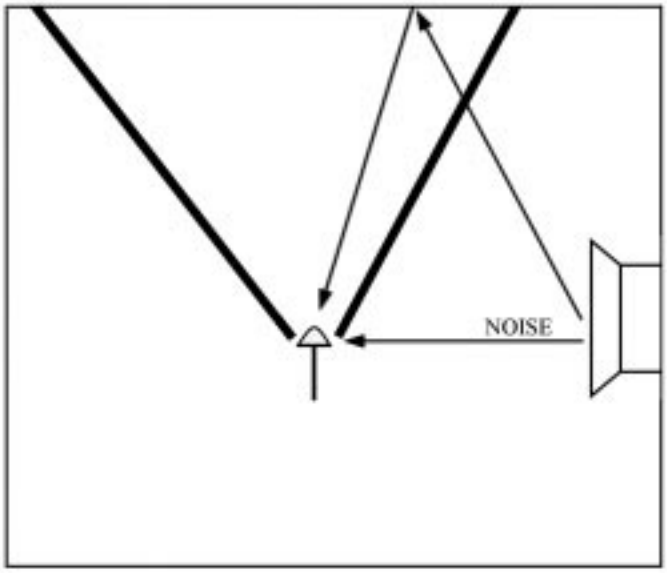

Figure 3c.
Reflected sound reaching a directional microphone (single reflection only).

Figure 3d shows two other more-complex sets of reflections. In one case, noise is reflected off the bottom wall and then again off the top wall. In another case, noise is reflected off the bottom wall, the left-hand wall, and then the top wall before reaching the microphone. As a result, the sound reaching the microphone will be coming from several directions.

It is still possible to reduce the amount of noise reaching the microphone by using a directional microphone, but the reduction in background noise is not nearly as great as when there are no reflections. This is because some of the reflected noise reaches the microphone from the same direction as the speech signal.

A room in which sound is reflected with little attenuation off walls, floors, ceilings, and other hard surfaces will result in an extremely large number of reflections reaching the microphone (or ear, for unaided listening). These multiple reflections are referred to as reverberation. Some reverberation is useful in reinforcing the

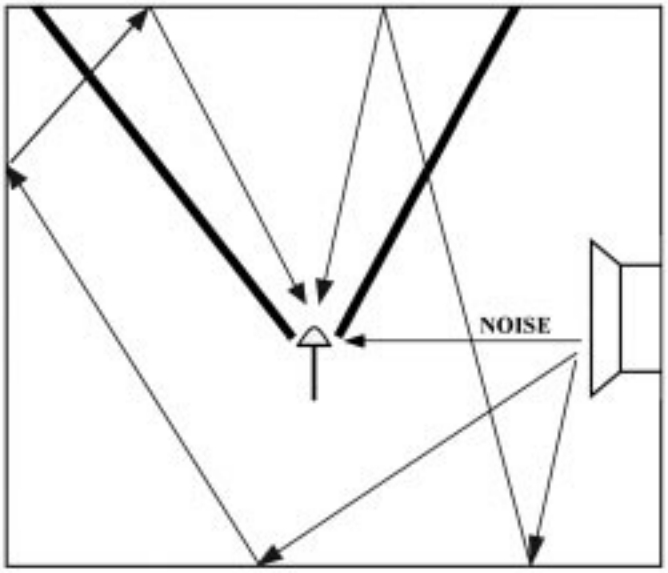

Figure 3d.
Reflected sound reaching a directional microphone (multiple reflections).

speech signal but excessive reverberation will reduce sound quality with a corresponding reduction in speech intelligibility.

The problem is more severe for hearing-aid users in that even moderate amounts of reverberation will not only reduce the quality of sound received in quiet but will also add significantly to the reduction in intelligibility of speech in noise. Further, as shown in **Figure 3d**, directional microphones are less effective in separating speech from noise under reverberant conditions (35).

It is possible to improve the directional properties of a hearing aid by using an array of microphones rather than a single microphone (36–38). Microphone arrays can be relatively small, including arrays that are compact enough to be used with a hearing aid that can fit on or in the ear. Slightly larger arrays that can be mounted on the stem of a pair of eyeglasses are capable of significantly greater directionality. The simplest microphone arrays add the signals received at each microphone after an appropriate delay so that the speech signal from each microphone is added in phase (i.e., the speech signal is increased in level by the maximum possible amount), while the noise is not added in phase (i.e., the noise is increased in level by an amount less than that of speech). Greater directionality is possible if, in addition to the delay, the signals are multiplied by a weighting coefficient before being summed. Since the direction of the speech and noise sources can vary over time, even greater

improvements can be obtained with an adaptive array that focuses on the speech source (39).

The potential gain in speech intelligibility using directional microphone arrays appears to be far greater than that which can be obtained from differences in the spectro-temporal characteristics of speech and noise. It should also be added that the signal processing used in implementing directional microphone arrays introduces far less processing noise than that produced by the computationally intensive techniques used for the more advanced methods of separating speech from noise in terms of spectro-temporal differences.

Recent advances in the microminiaturization of digital signal- processing chips have made the development of directional-microphone arrays for hearing aids a practical possibility. There is now an intensive research effort investigating possible hearing-aid applications of this technology. Issues being addressed include the development of greater directionality in microphone arrays of small size while also allowing some flexibility to reduce directionality as needed; e.g., being able to pick up warning signals from a direction other than that of the speech signal.

ADAPTIVE NOISE CANCELLATION

As noted in the Introduction, the more one knows about the speech and noise signals, the more effectively one can extract speech from noise. If, for example, the noise waveform is known exactly, then extracting the speech is a trivial problem. All that is necessary is to simply subtract the known noise waveform from the speech-plus-noise waveform and be left with speech only.

There are situations in which the noise waveform can be identified exactly. Consider the case of a single noise source in a typical room. It is possible to place a microphone at the location of the noise source so as to pick up noise only. A second microphone elsewhere in the room (e.g., on a hearing aid) will pick up both speech and noise. In order to subtract the noise from the speech plus noise picked up by the hearing-aid microphone, it is necessary to take into account the fact that there will be reflections of the noise off the walls of the room; i.e., by the time the noise gets to the hearing-aid microphone, the noise waveform will have changed.

It is possible to process the noise waveform so as to correct for these reflections. A special-purpose filter can be used for this purpose. If the filter is designed properly, subtracting the filtered noise from the speech plus noise picked up by the hearing-aid microphone will effectively cancel the noise with only the speech remaining (40). A system of this type is shown in **Figure 4**. Since people typically move around in a room, the pattern of reflections will change, and so it is necessary for the filter to keep adjusting itself. This method of noise reduction is known as adaptive noise cancellation.

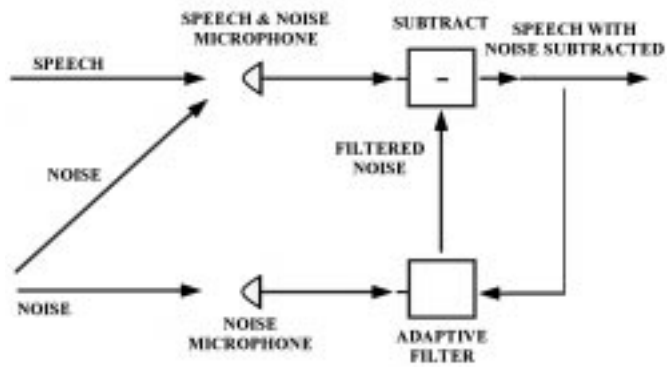

Figure 4.
Two-microphone adaptive noise cancellation.

Adaptive noise cancellation requires at least two microphones and, under ideal conditions, at least one microphone must be placed at the noise source. This is not very practical for a person wearing a hearing aid. It is possible, however, to have both microphones mounted on the head with one microphone picking up more speech than noise and the other microphone picking up more noise than speech, as shown in **Figure 5** (41,42). If the adaptive filter is used as before, the noise will not be cancelled completely but the level of the noise will be reduced. An improved speech-to-noise ratio will result, with improved intelligibility.

A variation of the above technique is to use a microphone mounted on each ear. Both the sum of, and the difference between, the two microphone signals are obtained. If speech is coming directly towards the listener and noise is coming from another direction, the sum of the two microphone signals will reinforce the speech signal, while subtracting the two microphone signals will cancel the speech, leaving the noise only. This situation is now equivalent to that of placing one of the microphones at the noise source to pick up noise only while the second microphone picks up both speech and noise. The adaptive noise- cancellation method is then used to cancel the noise (43).

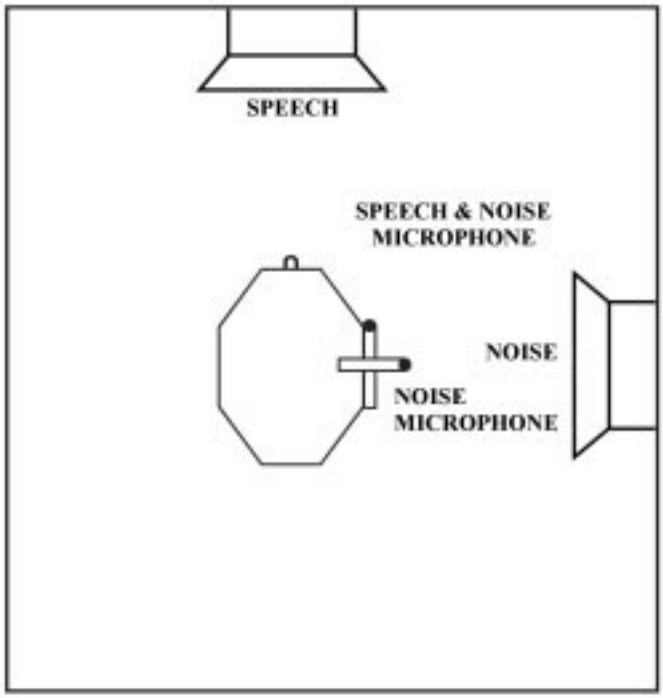

Figure 5.
Microphone configuration for a head-worn adaptive noise reduction system.

Hearing aids using adaptive noise cancellation are still at an early stage of development. Experimental evaluations of these techniques indicate that they share the strengths and weaknesses of directional microphone arrays. They work well when speech and noise come from different directions and there is relatively little room reverberation. They do not work well in a highly reverberant environment or with multiple noise sources. There is the possibility of combining elements of adaptive noise cancellation with directional microphone arrays to obtain improved performance. Even if these ongoing developments prove successful it will still be some time before practical hearing aids embodying this form of noise reduction become available.

KNOWLEDGE-BASED NOISE REDUCTION

Looking further into the future, it may be possible to use some aspects of speech-recognition technology to develop improved forms of noise reduction. Automatic speech-recognition devices make considerable use of our knowledge of the speech signal, including phonetic, linguistic, and statistical aspects of speech. Information of this type is not accessed using conventional methods of signal processing for noise reduction. It is conceivable that our knowledge of the non-acoustic properties of speech could be used to develop more effective methods of extracting speech from noise. For example, if an automatic speech-recognition device that has been trained on a specific talker is able to recognize the talker's speech in a moderately noisy environment, then the speech that has been recognized could be re-synthesized without any background noise. The speech recognition algorithms used in this process not only use information regarding the unique acoustic characteristics of that talker's speech (as obtained by training the speech-recognition device on a previous sample of the talker's speech), but also use a wealth of phonetic, linguistic, and statistical information drawn from our knowledge of speech.

Automatic speech recognition in noise is a very difficult problem, but given the recent rapid advances being made in this field it is not inconceivable that reasonably reliable automatic speech recognition in moderate amounts of background noise may be possible in the not-too-distant future. Noise reduction is similarly a very difficult problem, but there is a fundamental difference in current approaches to these two problems. Methods of extracting speech from noise have focused almost entirely on an acoustic analysis of the speech and the noise. The approach used in automatic speech recognition combines information obtained from an acoustic analysis with information extracted from a vast body of knowledge of both the acoustic and non-acoustic properties of speech.

It is significant to note that relatively little progress has been made in recent years in improving methods of extracting speech from noise using acoustic analysis only. It would appear that the limits of the purely acoustic approach have been reached. In contrast, substantial advances in automatic speech recognition have been made in recent years once large data bases became available which, in turn, allowed for the acquisition of considerable detailed information on the acoustic, phonetic, linguistic, and statistical properties of speech. Further, the best results in automatic speech recognition are obtained after the speech-recognition device has been trained on the speech of a specific talker; i.e., a detailed knowledge of the specific characteristics of a given talker's speech is of great value in automatic speech recognition. The point being made here is that future significant progress in extracting speech from noise is unlikely by restricting the analysis to the speech and noise signals only, without

drawing on the vast body of information now available on both the acoustic and non-acoustic properties of speech including, if possible, information on the specific characteristics of the talker whose speech is being partially masked by noise.

The speech-recognition speech-synthesis approach outlined above opens up new avenues of investigation not only with respect to the possibility of completely eliminating background noise (at least for moderate amounts of noise) but will also allow for the synthesis of speech, or reprocessing of speech and noise, so as to improve intelligibility by enhancing important spectral or temporal aspects of the speech signal that are not clearly perceived as a result of the hearing loss, even when listening in quiet (44).

Harry Levitt, PhD

REFERENCES

1. Gengel RW. Acceptable speech-to-noise ratios for aided speech discrimination by the hearing impaired. J Speech Hear Res 1971;11:219–22.
2. Plomp R. A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. J Speech Hear Res 1986;29:146–54.
3. Humes LE. Understanding the speech understanding problems of the hearing impaired. J Amer Acad Audiol 1991;2:59–69.
4. Levitt H, Neuman A, Mills R, Schwander T. A digital master hearing aid. J Rehabil Res Dev 1986;23(1):79–87.
5. Levitt H. New trends—Digital hearing aids: Past, present, and future. In: Practical Hearing Aid Selection and Fitting. Veterans Administration Rehabilitation Research and Development Service 1997; Monograph 001, p. xi-iii.
6. Van Trees HL. Detection, estimation and modulation theory. New York: Wiley, 1968.
7. Lim JS, Oppenheim AV. Enhancement and bandwidth compression of noisy speech. Proc IEEE 1979;67:1586–604.
8. CHABA. Removal of noise from noise-degraded speech signals. (Report of the) Committee on Hearing, Bioacoustics and Biomechanics, Commission on Behavioral and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press. 1989.
9. Dillon H, Lovegrove R. Single microphone noise reduction systems for hearing aids: A review and an evaluation. In: Studebaker GA, Hochberg I (eds.). Acoustical factors affecting hearing aid performance, second edition. Boston: Allyn and Bacon, 1993.
10. Weiss M, Newman AC. Noise reduction in hearing aids. In: Studebaker GA, Hochberg I (eds.). Acoustical factors affecting hearing aid performance, second edition. Boston: Allyn and Bacon, 1993.
11. Fletcher H. Auditory patterns. Rev Mod Phys 1940;12:47–65.
12. Patterson RD, Moore BCJ. Auditory filters and excitation patterns as representations of frequency resolution. In: Moore BCJ (ed.). Frequency selectivity in hearing. London: Academic Press, 1986.
13. Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. Hear Res 1990;47:103–38.
14. Greenwood DD. A cochlear frequency-position function for several species—29 years later. J Acoust Soc Am 1990;87(6):2592–605.
15. Bilger RC, Hirsh IJ. Masking of tones by noise. J Acoust Soc Am 1956;28:623–30.
16. Trees DE, Turner CW. Spread of masking in normal subjects and in subjects with high-frequency hearing loss. Audiology 1986;25:70–83.
17. Gagne J-P. Excess masking among listeners with a sensorineural hearing loss. J Acoust Soc Am 1988;83:2311–21
18. Shapiro I. Evaluation of the relationship between hearing threshold and loudness discomfort level in sensorineural hearing loss. J Speech Hear Dis 1979;64:31–36.
19. Levitt H. Hearing impairment and sensory aids. J Rehabil Res Dev 1986;23(1):xiii–viii.
20. American National Standards Institute. American National Standard Methods for Calculation of the Speech Intelligibility Index. ANSI S3.5, 1997.
21. Neuman AC, Schwander TJ. The effect of filtering on the intelligibility and quality of speech in noise. J Rehabil Res Dev 1987;24(4):127–34.
22. Ono H, Kanzaki J, Mizoi K. Clinical results of hearing aid with noise-level-controlled selective amplification. Audiology 1983;22:494–515.
23. Graupe D, Grosspietsch JK, Basseas SP. A single-microphone-based self-adaptive filter of noise from speech and its performance evaluation. J Rehabil Res Dev 1987;24(4):119–26.

24. Neuman AC, Eisenberg L. Evaluation of a dereverberation technique. J Commun Disord 1991;24:211–21.
25. Wiener N. Interpolation, extrapolation and smoothing of stationary time series. New York: J Wiley, 1949.
26. Levitt H, Bakke M, Kates J, Neuman A, Schwander T, Weiss M. Signal processing for hearing impairment. Scand Audiol 1993;Suppl 39:7–19.
27. Weiss M, Aschkenasy E. Automatic detection and enhancement of speech signals. Report No RADC-TR-75-77. Rome Air Development Center, Grifiss Air Force Base, NY, March 1975.
28. Boll SF. Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans Acoust Speech 1979;27(2):113–20.
29. McAulay RJ, Quatieri TF. Speech analysis/synthesis based on a sinusoidal representation. IEEE Trans Acoust Speech 1986;34(4):744–54.
30. Kates JM. Speech enhancement based on a sinusoidal model. J Speech Hear Res 1994;37(2):449–64.
31. Fabry DA, Leek MR, Walden BE, Cord M. Do adaptive frequency response (AFR) hearing aids reduce 'upward spread' of masking? J Rehabil Res Dev 1993;30(3):318–25.
32. Van Tasell DJ, Larsen SY, Fabry DA. Effects of an adaptive filter hearing aid on speech recognition in noise by hearing impaired subjects. Ear Hear 1988;9:15–21.
33. Tyler RS, Kuk FK. Consonant recognition evaluation of "Noise Suppression" hearing aids in speech-babble and low-frequency noise. Ear Hear 1989;10:243–9.
34. Van Tasell DJ, Thomas R, Crain MA. Noise reduction hearing aids: Release from masking and release from distortion. Ear Hear 1992;13:114–21.
35. Hawkins DB, Yacullo WS. Signal-to-noise ratio advantage of binaural hearing aids and directional microphones under different levels of reverberation. J Speech Hear Disord 1984;49:278–86.
36. Greenberg JE, Zurek PM. Evaluation of an adaptive beamforming method for hearing aids. J Acoust Soc Am 1992;91:1662–76.
37. Bilsen FA, Soede W, Berkhout AJ. Development and assessment of two fixed-array microphones for use with hearing aids. J Rehabil Res Dev 1993;30(1):73–81.
38. Soede W, Berkhout AJ, Bilsen FA. Development of a directional hearing instrument based on array technology. J Acoust Soc Am 1993;94:785–98.
39. Kates JM, Weiss MR. A comparison of hearing-aid array-processing techniques. J Acoust Soc Am 1996;99:3138–48.
40. Widrow B, Glover JR, McCool JM, Kaunitz J, Williams CS, Hearn RH, et al. Adaptive noise canceling: Principles and application. IEEE Trans Acoust Speech 1975;63:1692–716.
41. Schwander TJ, Levitt H. Effect of two-microphone noise reduction on speech recognition by normal hearing listeners. J Rehabil Res Dev 1987;24(4):87–92.
42. Weiss M. Use of an adaptive noise canceller as an input preprocessor for a hearing aid. J Rehabil Res Dev 1987;24(4):93–102.
43. Peterson PM, Durlach NI, Rabinowitz WM, Zurek PM. Multimicrophone adaptive beamforming for interference reduction in hearing aids. J Rehabil Res Dev 1987;24(4):103–10.
44. Levitt H. Advanced signal processing hearing aids. In: Beilin J, Jensen GR (eds.). Recent developments in hearing instrument technology, Proc 15th Danavox Symposium. Copenhagen: Stougaard Jensen, 1993.

ADDITIONAL READING

1. Lim JS (ed.). Speech enhancement. Englewood Cliffs, NJ: Prentice Hall, 1983.
2. Studebaker GA, Hochberg I. (eds.). Acoustical factors affecting hearing aid performance, second edition. Boston: Allyn and Bacon, 1993.
3. Levitt H. Processing of speech signals for physical and sensory disabilities. Proc Nat Acad Sci 1995;92:9999–10006.
4. Valente M, Hosford-Dunn H, Roeser M. Audiology treatment. New York: Thieme, 2000.