

Maximizing the Voice-to-Noise Ratio (VNR) via Voice Priority Processing

by Mark C. Flynn, PhD

The new digital aid uses parallel processing, adaptive directionality with separate polar patterns for frequency bands, as well as noise management and 8 channels of compression.

The processing of speech in noise to provide the optimum speech signal combined with appropriate noise management is one of the greatest challenges facing manufacturers of premium digital hearing instruments. The importance of this challenge is underlined by the fact that background noise remains the greatest problem reported by hearing aid wearers in addition to being a continuing source of dissatisfaction with hearing aids.^{1,2}

This article describes the Voice Priority Processing (VPP) system employed in the new Syncro hearing aid developed by Oticon. In this system, parallel processing is employed to drive the VPP system via Multi-band Adaptive Directionality (ie, separate polar patterns for each frequency band), TriState Noise Management (a system that categorizes noise situations into distinct listening strategies/modes), and Voice Aligned Compression (compression in eight independent channels across an expanded bandwidth). The overall purpose of the VPP system is to progressively optimize speech signals or, in essence, to help clean the signal of noise and maximize speech understanding.

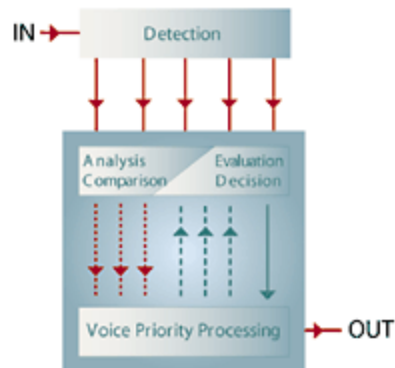
Artificial Intelligence: Sequential vs Parallel Processing

One of the challenges in designing advanced hearing instruments is to manage and select the best response of various systems, such as adaptive directionality or noise reduction. It is crucial that decisions and processing options are made quickly. If the group delay is greater than 10 ms, the user is likely to begin to hear echoes and other distortions.^{3,4}

This need for a fast processing time has heretofore prohibited the use of parallel processing—where all outcomes are evaluated for the best solution (eg, analogous to how advanced electronic chess games can look at all possible moves many turns into the future). Instead, hearing instruments have relied on sequential processing combined with a comprehensive prediction model to select the preferred processing option.^{5,6}

Because sequential processing does not actually compare different outcomes, an incorrect solution can be selected due to the unpredictability of complex communication environments. Parallel processing, however, provides the significant advantage of not relying on predictive models but instead processing and comparing each outcome for the best solution. In fact, parallel processing is a prerequisite for artificial intelligence that allows multiple processing schemes to be evaluated simultaneously to ensure that the best solution is implemented.

Figure 1. Stylized demonstration of the parallel processing between the Analysis & Decision cluster and the Voice Priority Processing system.









	Soft	Medium	Loud
Environment			
Available Dir Modes			

Figure 1 shows how the parallel processing interacts with the VPP system. Various acoustic detectors classify the acoustic environment and supply that information to the Analysis and Decision center. Information about the environment is sent to the VPP system for simultaneous

calculation of all processing options. The various calculations are then returned to the Analysis and Decision cluster for comparison of the possible outcomes to evaluate which one provides the best solution in terms of optimum Voice to Noise Ratio (VNR). VNR is a more apt description of the goal for this system compared to SNR (signal-to-noise ratio), because the focus is on the human voice rather than the most dominant signal (which, after all, could be noise). The decision is then returned to the VPP system for implementation, enabling the instrument to intuitively select the best solution, despite rapid changes and unpredictability of the listening environment.

Figure 2. Three different modes (Surround, Split and Full-Directionality) of the Multi-band Adaptive Directionality across different sound input levels. Surround uses the enhanced omni-directional microphones. Split-Directionality utilizes Surround in the low frequency band and full directionality in the remaining three high frequency bands. Full-Directionality implements full directionality in all four bands.

Artificial intelligence avoids the problem of trying to use complex models to predict the best response to the environment. The problem with prediction-based models is that they attempt to narrow the multiplicity of communication environments into a restrictive prediction-based formula. While the use of prediction works well in the laboratory, it does not always work as effectively in real-world communication environments; it's been shown that the prediction model may select incorrect settings as often as 30% of the time.⁵ In Syncro, the use of prediction models is avoided as underlying artificial intelligence (ie, parallel processing) is designed to analyze simultaneously all of the responses and select the best one—instead of using a restrictive prediction model.

Multi-band Adaptive Directionality

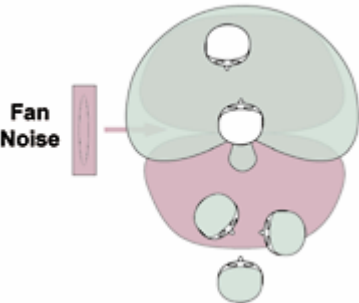
The only proven method to improve speech understanding in noise is through the use of a directional microphone.⁷ Directional microphones suppress the signal from the sides and/or rear of the wearer while retaining the audibility for signals from the front. The underlying assumption is that the user will be facing the person with whom they are talking and wish to perceive that signal in preference to other speakers or background noise. In this way the improvement in SNR has been estimated as around 3-4 dB SPL when speech and noise are presented in ways that are representative of real world listening environments.^{8,9}

One of the more recent advances in directionality has been the concept of adaptive directionality. Adaptive directionality overcomes two of the important drawbacks of a fixed directionality system. First, it should be recognized that directionality is not always the best configuration for all listening situations, especially when in the presence of reverberation or for listening at a distance.⁷ With adaptive directionality, the hearing instrument automatically decides whether to select an omni-directional or directional response. Second (and more impressive), recent adaptive directionality systems detect the angle of the dominant noise source and automatically change the polar pattern to provide maximum attenuation for that sound.¹⁰ Therefore, the null in the directivity pattern is always directed towards the most dominant noise source, ensuring a higher VNR than fixed directionality systems.¹¹

While adaptive directionality systems are impressive, they have four well-recognized limitations. First, due to increased microphone noise, they require a relatively large input level to switch to the directional mode. Second, directional microphones are susceptible to wind noise.^{10,12} Third, they cannot cancel multiple independent noise sources simultaneously,¹² because they only have one polar response across all frequencies. Fourth, the selection of the directional-vs-omnidirectional mode is based on a prediction of which mode will provide the best solution rather than making decisions based on which mode actually provides the best signal.

The Multi-band Directional Microphone system is a first-order directional algorithm based on inputs from two omni-directional microphones. The directional system has been expanded from a single-frequency algorithm¹³ to allow separate adaptive polar responses in each of the four frequency bands (termed Full-Directionality). Therefore, each frequency band could potentially have a different polar response.

Additionally, in the hybrid Split-Directionality mode, the low frequency band has an omni-directional signal while the three higher bands have separate adaptive directional signals. To ensure that the correct mode (Surround, Split- or Full-Directional) is selected, parallel processing evaluates simultaneously the various available polar responses in each frequency as well as the resulting VNR for each mode. The splitting of the directionality system into four separate bands combined with parallel processing of the signal is designed to solve the aforementioned problems associated with adaptive directional systems:



1. Directional benefit at lower sound intensity levels. While directional microphones can provide benefit in medium sound levels, automatic systems rarely switch to the directional mode as the sound environment cannot mask out the increased microphone noise. To increase the benefit of directionality in more sound environments, Multi-band Adaptive Directionality can be implemented in three modes with the unique Split-Directional mode providing an innovative solution to the problem of increased directional microphone noise.

- Surround Mode. When the sound input is soft, Surround is the best mode because there is little noise to attenuate. In this mode, the signal from the two omni-directional microphones is enhanced by adding the signals. This effectively reduces microphone noise by 3 dB, providing an improved signal for low input levels.
- Split-Directional Mode. For moderate inputs, the Multi-band Adaptive Directionality has a choice between Surround and the Split-Directionality mode. Traditionally, directional microphones need to be either omni-directional or directional in all frequencies. The system splits the microphone signal into four separate frequency bands. This allows the use of the enhanced Surround mode in only the lowest frequency band (below 1000 Hz) and maintains three independent bands of directionality in the mid and high frequencies. The use of the Surround signal in the lowest frequency band reduces directional microphone noise while providing directional benefit in the three other frequency bands. Split-Directionality is used when there is a medium-to-loud degree of background noise and the VNR judged as being best in that mode.
- Full-Directional Mode. Full-Directionality enables the ability for four bands of directionality with separate polar responses, providing the possibility to attenuate up to four separate noise sources simultaneously. Full-directional is implemented in loud environments where the VNR judged as being best in that mode.

2. Reducing wind noise. Airflow around the head generates noise within centimeters of our ears and also within the microphone ports.¹² This turbulence is especially a problem in directional microphones due to their increased sensitivity to sounds from the near field. Because of this, omni-directional microphones have been recommended in situations prone to wind noise,¹² and some adaptive directional systems will detect wind noise and turn off the directional microphone.¹⁰

Syncro provides an alternate solution for this problem, allowing the wearer to continue to benefit from directionality despite the presence of wind noise. The Split-Directionality mode is used in situations of moderate wind noise. In windy situations, the first band of the Multi-band Adaptive Directionality is set to the Surround mode so that the microphone is less sensitive to low-frequency wind noise. This decreases the annoyance of wind noise while allowing the user to benefit from three bands of directionality in the higher frequencies.

Figure 3. Automatic directionality determines the best VNR. A different polar pattern is used in the low frequencies to decrease the effect of a fan with low frequency energy while preserving directionality in the high frequency region.

3. Simultaneous attenuation of up to four independent noise sources. The configuration of the directional polar pattern is crucial to ensure the maximum amount of noise is attenuated. A problem of conventional

adaptive systems is that they can only create a single polar pattern for any one environment. Unfortunately, in complex listening situations, distracting noise arises from multiple sources and, while generally broadband, the noise may exhibit different spectral peaks. Noise sources with different amplitude peaks across frequencies provide an opportunity to separately target those signals. The four distinct frequency bands analyzed in this device enable the system to assign different polar responses for each band. Therefore, it is now possible to simultaneously cancel out up to four separate noise sources at one time. For instance, Figure 3 shows a situation in which there are two different noise sources behind the listener: a low frequency fan noise at 270° azimuth and three people talking at 180° azimuth. In this case, the hearing instrument will determine that the VNR is best if the lowest frequency band cancels the fan noise and the three higher bands focus on canceling the chatter from behind. The parallel processing in the system allows a simultaneous comparison of multiple polar plots to ensure that the one yielding the best VNR is used.

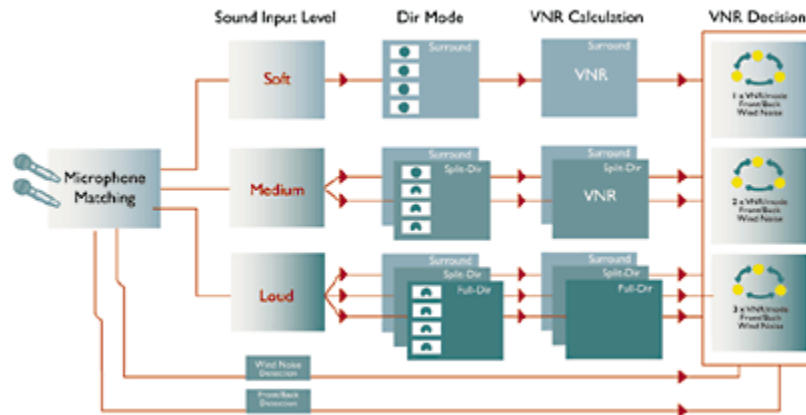


Figure 4. Multi-band Adaptive Directionality showing how the parallel processing tracks and the VNR decision process.

4. Decision-making through parallel processing. The purpose of the Multi-Band Adaptive Directionality system is to provide the best possible VNR to the user. The heart of this system is using parallel processing to achieve the best result (Figure 4). The input level (soft, medium, or loud) determines which modes of the Multi-band Adaptive Directionality should be made available. In each mode the best configuration of polar plots is then determined by evaluating the VNR in each combination (“Dir mode”). Once the optimum configuration of polar plots is determined, then the “VNR calculation” is performed. This calculation for each mode is sent to the “VNR decision” which decides which mode provides the best VNR. The wind and front/back detectors may overrule this decision if there is increased wind or the primary speaker is behind the listener. In this way, Multi-band Adaptive Directionality is designed to provide the best signal in terms of VNR to the TriState Noise Management System.

Underlying the artificial intelligence of the system is the ability to take information from various environmental detectors and analyze all the possible output combinations in order to yield the best VNR. This flexibility is designed to provide two advantages over former adaptive directional systems: 1) Split-Directionality offers the benefit of directionality at lower sound pressure levels and in the presence of wind (where other systems would be forced into the omni-directional mode); 2) the four bands of directionality (three in Split-Directionality) can attenuate up to four separate noise sources simultaneously.

TriState Noise Management

While noise management systems cannot improve speech understanding in noise, they do allow the background noise to be measured and channel-specific noise reduction to be implemented, improving listening comfort.^{14,15}

One of the challenges of noise management systems is to accurately describe the acoustic environment. Two approaches to solving this problem have been used: detecting speech¹⁶ or alternatively evaluating the amount of noise in the environment through modulation detection.¹⁴ A drawback of modulation-based noise management systems is that they have difficulty distinguishing speech from noise at poor VNRs.¹⁵⁻¹⁷ This

can lead to a compromise situation where the client can experience more noise reduction at the sacrifice of speech audibility, or have increased speech audibility at the sacrifice of comfort. This situation arises from the inability of modulation detection to correctly identify speech at poor VNRs, leading to the incorrect classification of that environment as being “noise without speech.” It is crucial that a noise management system can detect the presence of speech and apply different noise management techniques whether speech is present or absent in the noise. Studies such as those conducted by Souza & Kitch¹⁸ reinforced the need to ensure that attenuation is separately applied in noise management to avoid reducing the speech signal. Similarly, data-logging of actual communication situations showed that two-thirds of the time spent in noisy environments was spent listening to speech. This reinforces the need for the three separate states of noise management:

- Speech in Quiet to ensure optimum speech understanding across a wide range of input levels.
- Speech in Noise to limit the degree of channel specific noise reduction in accordance with the Articulation Index when in complex listening environments where speech is present so as to ensure that speech understanding is maintained.
- Noise Only. For optimal comfort in noise, the maximum attenuation in each channel is provided when there is only noise present in the signal.

The combination of the VoiceFinder speech detector¹⁶ with modulation detection allows the classification of these three environmental conditions. The VoiceFinder speech detector detects the high frequency synchronous energy, which is the unique property of the human voice. VoiceFinder is a speech detection system designed to accurately detect speech down to a VNR of 0 dB.¹⁶ Modulation detection, while having shortcomings in the detection of speech in challenging listening situations,¹⁵ is able to accurately determine the noise level and the degree of modulation—and hence the degree of masking and annoyance provided by background noise. The combination of these two systems is designed to provide accurate detection of speech and estimation of noise in the environment.

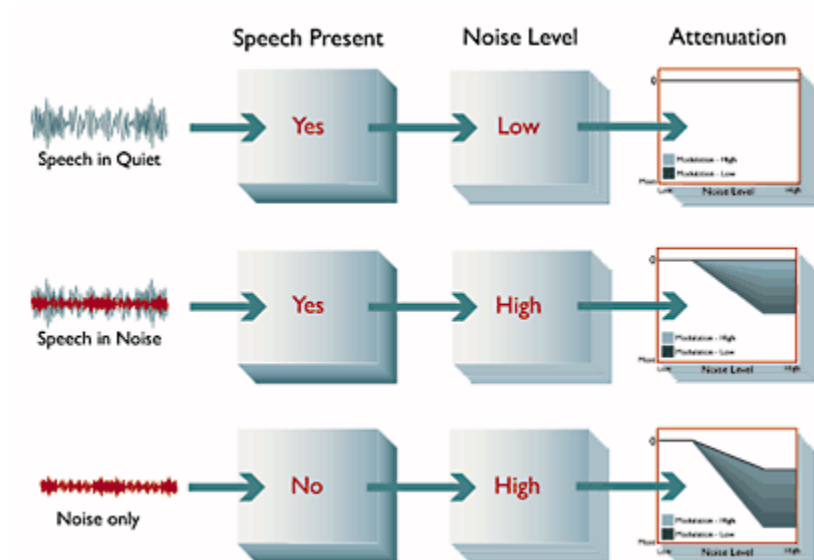
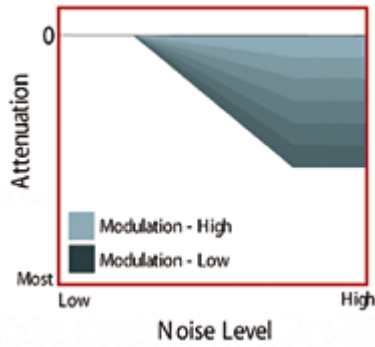


Figure 5. Schematic of the flow through the TriState Noise Management System.

The degree of noise reduction in both the Speech-in-Noise and Noise-Only modes is applied differentially, depending on the amount of actual noise level in each channel and the degree of modulation. Attenuation increases as a factor of noise level and of modulation index. The greater the noise and the lower the modulation, the higher the resulting attenuation. Importantly, the total amount of attenuation is less when speech is present versus when there is noise only in the signal. Additionally, when speech is present the noise reduction is shaped according to the Articulation Index to ensure speech understanding is maintained (Figure 5).

It is crucial for the Syncro Optimization Equation to optimize the input from the VoiceFinder Speech Detector, the Modulation Index, and the Noise Floor Level. Therefore, the speeds at which the TriState system switches between states are controlled to prioritize speech information and maintain listening comfort in difficult acoustic environments. The VoiceFinder speech detector has priority. Whenever speech is detected, the TriState system will move quickly into the Speech-Only or Speech-in-Noise settings. In contrast, a noise reduction system that quickly adapts to the presence of noise (through modulation detection only) may be over-reactive to transient noise signals, sacrificing comfort. Therefore, the reactions to noise—and in particular fluctuations in noise level—are more gradual in Syncro.



To manage effectively the transitions between speech-in-quiet, speech-in-noise, and noise-only, Syncro implements parallel processing where—irrespective of the environment—each of the three states of noise management are calculated simultaneously. This enables the system to better interpret and react to changes in the listening environment.

Figure 6. Demonstration of the attenuation in one state (Speech in Noise) of the Syncro TriState showing the progressive application of noise reduction as noise level increases and modulation decreases.

Through the use of the TriState Noise Management system, Syncro continuously analyzes the environment to provide the optimum amount of attenuation in any complex listening environment. When speech is present with background noise the degree of attenuation is dependent on the level and degree of modulation of the corrupting noise source (Figure 6). This provides the best speech understanding and comfort in each situation, without unnecessary suppression. When only noise is present, the TriState system is focused on comfort and will provide increasing noise reduction as the noise level increases and modulation decreases.

Compression Management

Voice Aligned Compression is built on a new platform of eight independent channels of compression across an expanded bandwidth. The amplification provides curvilinear compression comprising up to seven kneepoints to ensure a smooth frequency response at all input levels. This increase in channels and functionality provides increased opportunities to create and optimize the sound scheme for the individual user.

Compared to conventional amplification strategies, the Voice Aligned Compression provides less compression at high input levels (ie, amplifies to completion), and more compression at low input levels through a lower compression kneepoint (ie, increased gain for weaker inputs). Loudness compensation, per se, is not the main goal of this compression strategy; the goal is to provide improved subjective sound quality (“naturalness”) without loss of speech intelligibility. This objective is further enhanced by the progressive optimization of the signal provided by the Multi-band Adaptive Directionality and TriState Noise Management combined with the natural sound path from OpenEar Acoustics.

Previously, the purpose of WDRC systems was to place or squeeze speech within the reduced dynamic range of the listener.¹⁹ The assumption is that increased audibility of low input sounds will improve performance over linear amplification strategies. Looking at the research on compression, several authors report that, for listeners with mild to severe hearing impairment, compressed speech provides benefits at low and moderate input levels for both quiet and noisy listening situations.^{18,21} Similarly, it is interesting to note that, for speech understanding in quiet, the increase of compression ratio does not affect the understanding of speech.²¹ Therefore, it is clear that providing low level compression for low to moderate level inputs is beneficial. In these situations the listener is generally in a reasonable communication situation without a large degree of background noise or a poor VNR and can therefore tolerate a reasonable amount of compression.

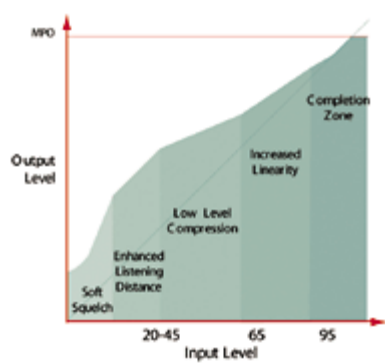


Figure 7. Input/Output function for the Voice Aligned Compression.

Unfortunately, while we can use compression to provide benefits at low and moderate input levels, high compression ratios in difficult listening situations do not provide benefit. In fact, both speech quality ratings and speech recognition scores decrease with increased compression ratios.²²⁻²⁴ For example, for speech presented in cafeteria noise (high input levels), people prefer low compression ratios (<1.5:1.0) or even linear compression.²³

Increasing the linearity for difficult listening situations provides greater access to the temporal cues in speech, which are more resilient to masking. Therefore, it is critical to provide less compression so as to maximize the access to temporal information. Additionally, communication situations at high intensity levels have low VNRs. Therefore, the short-term dynamic range of the environment is reduced, and compression is not needed to provide audibility of low-level signal portions. In these situations, improved communication is facilitated by reducing the negative side effects of compression.

The key benefit of the Voice Aligned Compression is the ability to provide both low-level compression, as well as increased linearity for high-level signals. Low-level compression combined with increased listening distance ensures audibility of soft speech sounds and improves communication at a distance. Whereas, increased linearity ensures that, in difficult listening situations, the side effects of compression are reduced and more temporal cues are available to assist in speech understanding.

Progressive Optimization of the Signal

While each of the VPP systems process signals in parallel to determine the best state, the signal is being progressively optimized to provide the optimal processing of the speech signal (Figure 8). First, the Multi-band Adaptive Directionality selects the best overall VNR. The TriState Noise Management provides state-specific attenuation of noise on a channel-by-channel basis, and delivers the best possible signal to the Voice Aligned Compression. As discussed, the signal is constantly being improved and progressively optimized. As opposed to a simple measurement of the input and a static set of adjustments, the ongoing state of the signal is being monitored and changes are made as needed.

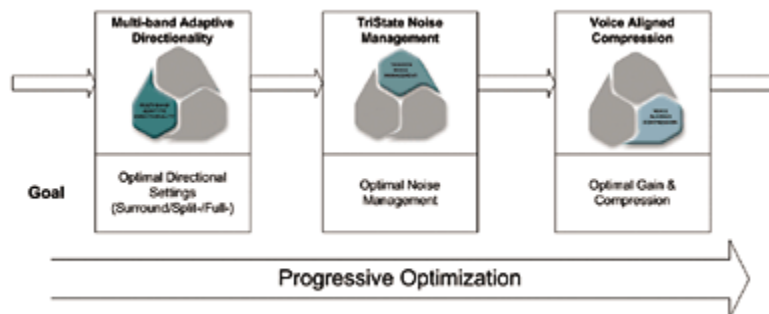


Figure 8. Progressive Optimization of the signal ensures that the three core components of VPP work together to assure the best possible VNR.

Summary

The goal of Voice Priority Processing is to provide the best possible VNR to the person with a hearing impairment through optimal amplification of speech and management of noise. This is achieved by the combination of Multi-band Adaptive Directionality, TriState Noise Reduction, and Voice Aligned Compression. Underlying these three systems is the decision-making provided by parallel processing, which enables every possible configuration of the Syncro's systems to be evaluated prior to implementation so that the best solution can be selected. Combining the systems to work in synergy allows the signal to be progressively optimized so that speech understanding is prioritized.

References

1. Kochkin S. MarkeTrak VI: 10-year customer satisfaction trends in the US hearing instrument market. Hearing Review. 2002; 9(10):14-25,46.
2. Kochkin S. MarkeTrak VI: Consumers rate improvements sought in hearing instruments. Hearing Review.

2002; 9(11):18-22.

3. Stone MA, Moore BCJ. Tolerable hearing aid delays I. Estimation of limited imposed by the auditory path alone using simulated hearing losses. *Ear Hear.* 1999; 20:82-192.
4. Stone MA, Moore BCJ. Tolerable hearing aid delays. II. Estimation of limits imposed during speech production. *Ear Hear.* 2002; 23(4):325-338.
5. Gabriel B. Study measures user benefit of two modern hearing aid features. *Hear Jour.* 2002; 55(5):46-50.
6. Kates JM. Classification of background noises for hearing aid applications. *J Acoust Soc Am.* 1995; 97:461-470.
7. Walden BE, Surr RK, Cord MT. Real-world performance of directional microphone hearing aids. *Hear Jour.* 2003; 56(11):40-47.
8. Amlani AM. Efficacy of directional microphone hearing aids: a meta-analytic perspective. *J Am Acad Audiol.* 2001; 12(4): 202-214.
9. Wouters J, Litiere L, van Wieringen A. Speech intelligibility in noisy environments with one- and two-microphone hearing aids. *Audiology.* 1999; 38(2):91-98.
10. Valente M, Mispagel KM. Performance of an automatic adaptive dual-microphone ITC digital hearing aid. *Hearing Review.* 2004; 11(2):42-46,71.
11. Ricketts T, Henry P. Evaluation of an adaptive, directional-microphone hearing aid. *Int J Audiol.* 2002; 41(2):100-112.
12. Thompson SC. Tutorial on microphone technologies for directional hearing aids. *Hear Jour.* 2003; 56(11):14-21.
13. Elko GW, Pong AN. A steerable and variable first-order directional microphone. In: *Proceedings of the IEEE ICASSP Workshop; April 1997.*
14. Alcantara JI, Moore AP, Kuhnel V, Launer S. Evaluation of a noise reduction system in a commercial digital hearing aid. *International J Audiol.* 2003; 42:34-42.
15. Schum D. Noise-reduction circuitry in hearing aids: (2) Goals and current strategies. *HearJour.* 2003; 56(6):32-41.
16. Elberling C. About the VoiceFinder. *News From Oticon: Audiological Research Documentation.* 2002; 3:1-11.
17. Flynn MC. Maximizing speech understanding and listening comfort in noise. *Hearing Review.* 2003; 7: 50-53.
18. Souza PE, Kitch VJ. Effect of preferred volume setting on speech audibility in different hearing aid circuits. *J Am Acad Audiol.* 2001; 12(8):415-422.
19. Dillon H. Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times? *Ear Hear.* 1996; 17(4), 287-307.
20. Kim ACS, Wong LLN. Comparison of performance with wide dynamic range compression and linear amplification. *J Am Acad Audiol.* 1999; 10, 445-457.
21. Souza PE, Turner C. Quantifying the contribution of audibility to recognition of compression-amplified speech. *Ear Hear.* 1999; 20:12-20.
22. Boike KT, Souza PE. Effect of compression ratio on speech recognition and speech quality ratings with wide dynamic range compression amplification. *J Sp Lang Hear Res.* 2000; 43: 456-468.
23. Neuman A C, Bakke MH, Hellman S, Levitt H. Effect of compression ratio in a slow acting compression hearing aid: Paired comparison judgements of quality. *J Acoust Soc Am.* 1994; 96(3):1471-1478.
24. Neuman A C, Bakke MH, Mackersie C, Hellman S, Levitt H. The effect of compression ratio and release time on the categorical rating of sound quality. *J Acoust Soc Am.* 1998; 103(5):2273-2281.