informa
healthcare

Original Article

International Journal of Audiology 2010; 49: 434–443

Assessing the benefit of adaptive null-steering using real-world signals

William S. Woods
Ivo Merks
Tao Zhang
Kelly Fitz
Brent Edwards

Starkey Laboratories, Inc., Eden Prairie, Minnesota, USA

Key Words

Hearing aid
Directionality
Beam-forming
Null steering

Abbreviations

AIDI: Articulation-index weighted directivity index
BTE: Behind-the-ear
DRR: Direct-to-reverberant energy ratio
MVDR: Minimum variance distortionless response
SII: Speech intelligibility index
SNR: Signal-to-noise ratio

Abstract

This study compared the noise reduction of adaptive null-steering and near-hypercardioid directional hearing-aid algorithms via performance on real-world signals. Using subject-individualized and generic (i.e. similar to current hearing aids), off-line frequency-domain implementations, we processed recordings made through two microphones of a BTE device worn by five subjects. Recording scenarios included homes, offices, cafés, streets, buses, and automobiles. We found practically all (> 95% of recording time) adaptive noise-reduction benefit for generic implementations is below 1.2 dB, and 96% and 92% is below 2 dB for 16- and 32-band individualized implementations, respectively. A 256-band, individualized implementation showed a majority of benefit between 1–4 dB. We found no extended (> 2 s) continuous periods of significant (> 2 dB) benefit for the generic adaptive implementations. The recordings—having many independent and simultaneously active sources, spatially extended sources, significant reverberation, or combinations thereof—indicate an environment comprising few instances of high direct-to-diffuse energy situations. Combined with results from previous field trials, the evidence suggests that such an environment is common and represents a significant limitation on adaptive benefit.

Sumario

Este estudio compara la reducción del ruido en algoritmos de auxiliares auditivos con direccionamento anulado o direccionamiento casi-hipercardioides por medio del rendimiento con señales de mundo real. Usando implementaciones fuera de línea en campos frecuenciales, individualizadas y genéricas (p.ej. similares a las de los auxiliares auditivos usuales), procesamos los registros realizados con dos micrófonos de un instrumento BTE usado por cinco sujetos. Los escenarios de registro incluyeron hogares, oficinas, cafés, calles, autobuses y automóviles. Encontramos prácticamente en todas las implementaciones (> 95% del tiempo de registro) que el beneficio en la reducción de ruido adaptable es inferior a 1.2 dB y que en 96% y 92% es menor a 2 dB para implementaciones individualizadas en bandas de 16 y 32, respectivamente. Una implementación individualizada en una banda de 256, mostró la mayoría de los beneficios entre 1-4 dB. No encontramos períodos continuos extendidos (> 2 s) de beneficio significativo (> 2 dB) con las implementaciones genéricas adaptables. Los registros –que tienen muchas fuentes independientes o simultáneamente activas, extendidas espacialmente, con reverberación significativa o combinaciones de los mismos- indican un ambiente que comprende pocas instancias de situaciones de alta energía de difusión directa. La evidencia, combinada con los resultados de pruebas libres previas, sugiere que este ambiente es común y representa una limitación significativa para el beneficio de adaptación.

Much work has been done to characterize the benefit of directional hearing aid processing in the laboratory and field. While a complete description of all work is beyond the scope of this paper, the work of Walden and colleagues (Cord et al, 2002, 2004; Walden et al, 2000, 2003, 2004, 2005) provides a systematic and representative comparison between fixed directional and omnidirectional processing. In early work (Walden et al, 2000) they showed that, despite significant directional benefit in the laboratory (based on measurements of speech intelligibility in noise), no significant perceived benefit could be found in the field (based on subjective ratings). In a series of subsequent investigations (summarized in Walden et al, 2005) they found that environmental factors played a major role in the lack of significant perceived benefit. Essentially, signal, noise, and acoustic characteristics combined to yield relatively few situations in which the fixed directional processing benefit could be realized. That is, fixed directional processing does provide benefit over omnidirectional processing in real world scenarios, but these types of scenario are not common enough to yield a significant overall subjective benefit. In addition, Cord et al (2002) provide data indicating that omnidirectional is actually preferred over fixed directional for some situations.

We are now seeing a similar characterization of directional benefit arising in the context of *adaptive* versus *fixed* directional algorithms. For instance, in a recent clinical trial of a hearing aid with second-order adaptive directional processing, Palmer et al (2006) found their data suggested that adaptive processing '...does not create a listening condition that is readily differentiated from, nor preferred over, the fixed-directional (hypercardioid) or omnidirectional listening conditions.' (p. 199). A similar conclusion was reached by Bentler et al (2004). These reports are based on extensive field testing: Palmer et al had 49 subjects (27–85 years of age) evaluating combinations of omni-, static- (hypercardioid), and adaptive-directional implementations over a total of six days; Bentler et al had 10 subjects (42–80 years of age) evaluate (generic) static- (hypercardioid) and adaptive-directional implementations separately over two, three-week periods. We note, also, that these field trials may have been (uncontrollably) biased toward outcomes favoring the adaptive systems since: (1) Palmer et al (2006) used an aid that only allowed comparison of a static hypercardioid (a fixed hypercardioid is optimal under diffuse conditions, or conditions where noise-source direction is assumed to be uniformly probable from any direction around a listener) against a 2nd-order adaptive system, and (2) Bentler et al (2004) used an aid that only allowed comparison of a static cardioid against a 1st-order adaptive system. Given the extent of the trials and the potential bias, it is significant that no perceptible difference between adaptive and fixed systems was found.

That these field trial results may, again, be due to environmental factors is supported by work presented in Woods and Trine (2004). They argued that the theoretical benefit of adaptive null-steering over the average performance of the best fixed algorithm was

ISSN 1499-2027 print/ISSN 1708-8186 online
DOI: 10.3109/14992020903518128

Received:
February 3, 2009
Accepted:
November 25, 2009

William S. Woods
Starkey Hearing Research Center, 2150 Shattuck Ave.
#408, Berkeley, CA 94704, USA.
E-mail: william_woods@starkey.com

only 2.0 dB in the most favorable of expected real-world conditions. When combined with Killion's (2004) argument that, while relatively small increases in SNR may yield better hearing in noise, '... an AI-DI as low as 2 dB is unlikely to be noticed by the patient in a real-world setting.' (p. 16), the Woods and Trine (2004) result is consistent with environmental conditions contributing to limited perceived benefit.

The only constraint Woods and Trine (2004) assumed on adaptive benefit was the direct-to-reverberant energy ratio (DRR) of the noise scenarios considered. Adaptive algorithms would have greatest advantage in high DRR scenarios, since they could steer a null toward the direct noise source and essentially cancel noise at the directional output. As the DRR decreases, however, the adaptive algorithm's best option is to steer the null toward a hypercardioid pattern (the reverberant energy was assumed to be diffuse), which is the pattern yielding the greatest attenuation of a diffuse field (see Ricketts and Hornsby, 2003; and Kates, 2008, for further discussion of the effect of DRR on directional outcomes). The comparison fixed algorithm had a pattern that attenuated a diffuse field by nearly the same amount as a hypercardioid, yielding insignificant adaptive benefit in low DRR scenarios. In addition, based on assumptions concerning room size, absorption coefficients, and distance to a single jammer, the DRR of real-world conditions was estimated to be less than 5.0 dB. This all combined to yield the maximum expected benefit of adaptive over fixed to be 2.0 dB, and a 2.0 dB benefit in only some conditions may not be significant enough or occur often enough to be discernible to subjects in the field.

While the Woods and Trine (2004) results are consistent with the Palmer et al (2006) and Bentler et al (2004) results, there are still open questions regarding the conclusion of primary effect of environment. For example, Woods and Trine's assumptions regarding range and occurrence of DRR values in the real-world may be incomplete. In addition, other processing in the hearing aid in the clinical trials (e.g. digital noise reduction) may have influenced results. Also, the only actual data regarding environments visited in the field trials is subjective. That is, no real measurements of the environment characteristics or the aid performance in those environments have been carried out. These questions motivate the main goal of the current study, viz., to compare quantitatively the performance of different directional algorithms with real-world noise signals. Such an analysis furthers our understanding of benefit provided by current and future adaptive directional technology, which is required for an informed cost/benefit analysis.

One cost of adaptive directionality is related to its performance under channel mismatch. Essentially, while channel mismatch will not affect the ability to steer a null *per se*, it does admit the possibility of steering a null toward the target, resulting in target attenuation. Channel mismatch may arise due to changes over time between the responses of the microphones in a directional array, which are ultimately due to differential changes in the physical characteristics of the diaphragms in the microphones with exposure to environmental conditions. This is not an issue for currently-available, single-diaphragm static designs.

Other sources of channel mismatch arise from differential accumulation of debris in microphone ports, and from designing the array in one context (e.g. in free space) and employing it in another (e.g. on a head), which we refer to as design/application mismatch. Designs of static and adaptive systems tend to compromise performance in an attempt to mitigate the effect of design/application mismatch. Since data on the frequency and effect of such mismatch are not readily available, however, we provide analysis here of both compromised and uncompromised systems, in order to better understand the tradeoffs involved.

Another cost of adaptive directionality is the potential for attenuating target reverberation that aids speech understanding. The relative effects of early and late target reflections have been studied for decades, and a recent analysis (Hodgson & Nosal, 2002) and empirical data (Yang & Bradley, 2009) emphasize that the relative benefit or detriment of the reflections varies with changes in spatial conditions. If, however, an early target reflection is strong enough an adaptive system may steer a null toward it and attenuate it more (on average) than would a static system. Our design—which precluded recording with a target source present—did not allow us to quantify this potential adaptive drawback (however small it may be) and thus it may overestimate the adaptive benefit. A related contribution of the lack of targets to overestimation of benefit arises because our implementations' adaptation is not influenced by the presence of target reflections. In this case the attenuation of non-target sources is not decreased by steering away towards a target reflection. Current hearing-aid systems do not prevent the influence of target reflections on adaptation and thus may not yield as much benefit as found here.

Methods

The goal of this study was accomplished by laboratory processing of two-channel recordings made in real-world environments. These recordings were made on different normal-hearing individuals as they went about everyday activities wearing a digital recorder connected to the outputs of two microphones housed in a BTE hearing aid case in an endfire configuration. Floating-point processing on a digital computer in the laboratory allowed complete control over the type of, and constraints on, directional algorithm implementations.

Hardware and calibration

All recordings were made with a two-channel Edirol R09 digital recorder, sampling the outputs of a two-microphone BTE at 44.1 kHz using 24-bit samples. The two microphones were standard BTE microphones in an endfire configuration, and were current-buffered before output to the recorder. The output bandwidth was limited to 7.5 kHz by the buffering hardware. The BTE was connected to the recorder via a cable which was clipped to the subject's clothing with enough length to allow a non-encumbering range of motion. The recorder was carried in a pocket of the subject's clothing. Recordings were transferred digitally from the recorder to a computer where they were downsampled to 16 kHz before any further processing.

Before each use in the field, laboratory recordings were made of a computer-generated stimulus output by a Tannoy System 600 loudspeaker at 90° with the BTE in free space, and also on the subject from 0°, 45°, ..., 360°. The loudspeaker and aid (in free space and when on the subject) were 1.15 m off the floor. When in free space the aid was 0.5 m in front of the loudspeaker. The subject sat facing the loudspeaker with 1 m between the loudspeaker and the center of the subject's head. The free-space recording was used to match the two channels of the system, and the on-subject recordings were used to derive target-direction information for some of the implementations and also to characterize performance on each subject. Recordings in free space from 0°, 115°, and 180° were also made for use in implementing algorithms. The laboratory length, width, and height were 4.5 m, 3.8 m, and 3.0 m, respectively, and the aid was always at least 1.9 m from the closest wall. The floor is carpet on concrete; the walls are gypsum board partially covered with foam pads (Echo Eliminator Bonded Acoustical Pad, from Acoustical Surfaces, Inc.); and the drop-ceiling is acoustic tile. All laboratory recordings were made with Golay-code stimuli (Foster, 1986) which allowed determination of the impulse response from digital source to digital recording. Anechoic designs were then executed through the use of the rectangularly-windowed first 5 ms of the impulse responses, eliminating later reflections from any influence[1].

Individuals and environments

One female and four males participated in the recording. They ranged in height from 1.6 to 1.9 m, representing a significant range in body and head size (anthropometric features that might significantly influence acoustics at the two microphones). All wore the BTE on their right ear. Recordings were made in offices, homes, urban and suburban streets, cars, buses, trams, stores, restaurants, and cafés. Subjects were asked to go about their normal activities, recording for at least half a minute in each new scenario encountered. If the scenario were more 'dynamic' they were asked to record for a longer time.

In order to simplify evaluation, no 'targets' were present during the recordings. To accomplish this we asked the subjects not to record while engaging in face-to-face conversation. In this way processing is characterized by simply comparing the levels of the directional output and front channel. By sampling for just long enough to characterize a given scenario, subjects did not have to avoid situations requiring their conversing. Having no targets present mimics systems that do not adapt when a target is present (cf. Greenberg & Zurek, 1992). Performance for such systems is better than those that do adapt when targets are present since, even if they do employ constraints to avoid target cancellation (e.g. by not allowing nulls to steer into the front hemisphere), such systems incorporate target information into the null steering and null steering is optimally done with noise information alone.

The recordings were categorized into three environments: 'interior', 'exterior', and 'transport'. This allowed accumulation of performance data over time without 'smearing' information across environments that were patently different from one another. Total recording time was 112, 53, and 48 minutes in the interior, exterior, and transport environments, respectively. Duration in each environment is broken down in more detail in Table 1. Durations of recordings in different scenarios range from 0.3 to 15.5 min.

Directional processing

Directional algorithms were implemented in the frequency domain using linear filter banks covering a 0–8 kHz range with a spacing of 500, 250, or 31.25 kHz between bands (yielding 16-, 32-, and 256-band implementations). Filtering and output waveform synthesis were implemented using standard windowing/FFT overlap-add techniques (Allen, 1977; Crochiere & Rabiner, 1983). Analysis-window lengths were 64, 128, and 512 samples in the 16-, 32-, and 256-band cases, respectively. The number of samples between analyses was 8, 16, and 128 samples in the 16-, 32-, and 256-band cases, respectively. Channel matching was performed in the frequency domain of the front channel before any other processing. This was done by multiplying the signal in each band by a different complex-valued coefficient. A different set of coefficients was used with each different frequency analysis. The delay imposed by the analysis/synthesis process was 2.5, 5.0, and 24 ms for the 16, 32-, and 256-band systems, respectively. The 16- and 32-band systems provide frequency resolution similar to that currently available in marketed adaptive aids. The 256-band system was run to provide data on the potential benefit of increased resolution (of course, a real-world implementation would need to have a lower delay; cf. Stone & Moore, 1999, 2002, 2003, 2005).

Two adaptive/static pairs of algorithms were implemented, one pair optimized to each individual (to evaluate ideal performance) and another pair designed to operate in free space (i.e. with no objects present; done to evaluate performance of implementations more closely related to those in current hearing aids). These shall be referred to as 'individualized' and 'generic' implementations, respectively. All these algorithms operate in basically the same manner: apply complex-valued weights (comprising a gain and a phase delay) to front and rear channel signals, and then add the weighted signals. The difference in the algorithms lies in how the weights are designed. Details of the designs are given in the following paragraphs and summarized in Table 2.

Weights for the individualized adaptive/static pair were computed using the minimum variance distortionless response beamformer (MVDR; Cox et al, 1987; Kates & Weiss, 1996; Lockwood et al, 2004). The MVDR algorithm uses target-direction information (the 'steering vector') and the auto- and cross-correlation of the front and rear channel signals to minimize the directional output power while maintaining unity gain in the target direction. Steering vectors were derived from the individual recordings at 0° described above. The adaptive MVDR implementation smoothed the correlation values with a 50-ms time constant (as used in, e.g. Greenberg & Zurek, 1992)[2]. The correlation values of the individualized static MVDR implementation were fixed at the values determined by averaging over the complete set of recordings for a given individual.

Weights for the generic adaptive/static pair were based on placing nulls at certain angles in response to signals in free space. The generic adaptive implementation was based on the adaptive null steerer of Elko and Pong (1995). This implementation minimizes output power

Table 1. Breakdown of total duration in each environment in minutes. The 'Dry' and 'Rev.' classifications are based on the level of reverberation apparent in the recordings, with the former being a reverberation typical of a small building foyer, and the latter a higher reverberation typical of a large, hard-surfaced hall.

	Interior						*Transport*			*Exterior*	
	Home	*Office*	*Dry Bldg.*	*Rev. Bldg.*	*Shop*	*Restaurant/Café*	*Bus*	*Tram*	*Car*	*Residential*	*Urban*
Duration (minutes)	39.1	19.3	13.2	4.8	18.0	18.0	9.2	31.9	6.7	9.2	44.2

Table 2. Summary of algorithm type and design. 'R' refers to the correlation matrix comprising auto- and cross-correlation of front and rear channels.

	Algorithm	*Weight computation*	*Design*
Individualized	Static	MVDR; R fixed at long-term average.	On head
	Adaptive	MVDR; R computed online and smoothed with 50-ms time constant.	On head
Generic	Static	MVDR; R fixed at value for anechoic noise source at 115°	In free space
	Adaptive	Elko-Pong (1995); Adaptively combine front- and rear-facing cardioid outputs	In free space

by adaptively weighting the output of a rear-facing cardioid (null at 0°) and subtracting it from the output of a front-facing cardioid (null at 180°). The cardioids are designed using the free-space recordings described above, and the adaptive weight is limited to provide nulls only in the rear hemisphere. The adaptation rate was chosen to match that of the adaptive MVDR implementation. The weights of the generic static implementation were designed to yield a null in free space at approximately 115° (a compromise between maximum diffuse attenuation—with a null at approximately 110°—and lowered sensitivity to implementation calibration errors).

Quantifying performance

Articulation-index-based methods are commonly used to quantify directional performance (cf. Greenberg & Zurek, 1992, pp. 1666–1667). Because we are working with time-varying processing, we used a time-varying articulation index (Rhebergen & Versfeld, 2005). The full method is outlined in Figure 1. The signals from the front channel and directional output were first split into the 1/3-octave bands of the Speech Intelligibility Index (SII; ANSI, 1997) using linear-phase 2048-point FIR filters. The filter outputs were then squared, grouped into blocks of samples ('frames') with frame duration as used in Rhebergen and Versfeld (2005), and summed within frames. This yielded signals (one per band) proportional to the time-varying power in each filter band, sampled at a rate of approximately 106 kHz (i.e. the inverse of Rhebergen and Versfeld's frame rate of 9.4 ms; frame size varied inversely with band center frequency from 35 ms in the lowest-frequency band to 9.4 ms in the highest-frequency band). Signal values were then converted to decibels, and the difference between front-channel and directional-output paths was weighted by the average speech weights (as proposed by Pavlovic (1987), for daily communication activities) in each band, and the result summed across bands. This yielded the time-varying 'directional' gain.

Because the implementations did not necessarily provide unity gain in the target direction[3], a moment-by-moment target gain was computed and subtracted from the directional gain, forming the 'system' gain. Target gain was determined by applying the directional weights as computed by the directional algorithms (which were either fixed or adapting in response to the recorded environmental signals) to a simulated anechoic target (cf. Greenberg & Zurek, 1992, pp. 1666–1667). Simulated targets were generated for each individual by convolving a white Gaussian noise with the individual's 5-ms windowed, laboratory-measured target-direction impulse response for the front and rear channels. The front channel and directional output in response to the simulated target was processed according to the articulation-based method described in the previous paragraph. The output of this process is defined to be the target gain. Unless otherwise indicated, in the remainder of this work all references to 'gain' calculated here are to the system gain, that is, the difference between directional and target gains. Note that, as when viewing polar patterns showing directional performance, lower values of gain indicate greater noise reduction.

Results

Results are first examined as two-dimensional histograms of adaptive/static system gain pairs. Figure 2 is a block diagram of the method used to determine the histogram data. The front and rear channel recordings for a given subject/environment are fed through the static and adaptive directional implementations and gain determination in parallel. Each time-aligned pair of frames of the gain

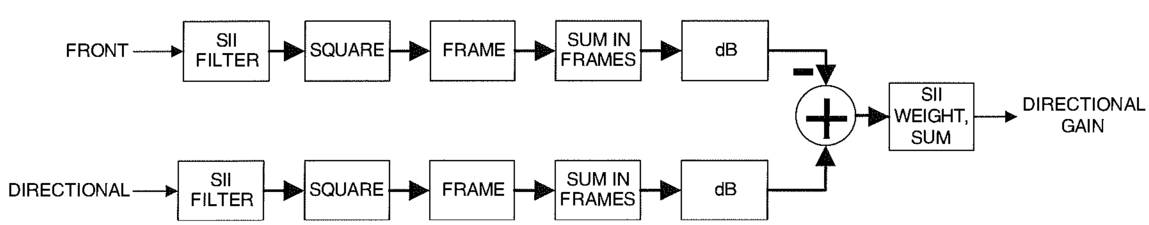

Figure 1. Block diagram of directional gain computation. The front-channel signal FRONT and directional output signal DIRECTIONAL are used to compute the directional gain. Thin arrows represent time-domain signals and thick arrows represent parallel frequency-domain signals. The input signals were first split into the 1/3-octave bands of the SII (ANSI, 1997) using linear-phase 2048-point FIR filters. The filter outputs were then squared, grouped into frames of samples with frame-duration as used in Rhebergen and Versfeld (2005), and summed within frames to obtain values proportional to power in each band. The summed values were then converted to decibels, and the difference between upper and lower paths was weighted by the average speech weights in each band, and the result summed across bands.

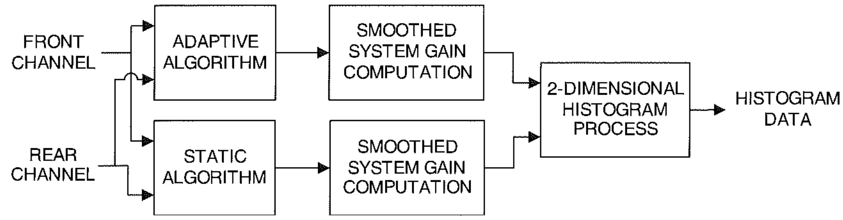

Figure 2. Block diagram of histogram method. The front and rear channel signals are processed by adaptive and static directional algorithms in parallel, followed by parallel gain computations. This allows incrementing histogram bin counts based on time-aligned pairs of adaptive/static gain values.

signals is used to increment the count in the appropriate histogram bin, and the bin data are saved for plotting.

Figures 3 and 4 show histograms of the moment-by-moment gains of the individualized and generic implementations, respectively, accumulated across subjects and presented as contour plots. Each panel shows results for a different combination of number of bands and environment. The contour plots tend to be elongated along a diagonal direction, with the majority of the data falling on or between the two diagonal lines (note that the grayscale is logarithmic—the value in the lower right of each panel shows the proportion of data between the two lines). The right diagonal line shows gain pairs yielding adaptive gain equal to static gain (i.e. 0 dB adaptive benefit), and the left line shows gain pairs with adaptive gain 2 dB lower than static gain (i.e. 2-dB adaptive benefit). Thus, the contour plots indicate an adaptive benefit for the 16- and 32-band individualized implementations and all generic implementations lying mainly between 0 and 2 dB. The benefit for the best-performing implementation is not much greater than this: The majority of data for the individualized, 256-band implementation lies between 1–4 dB of benefit.

For the individualized case (Figure 3) the histograms tend to have data farther to the left of the diagonal at the lower left and the upper right of the panels than in the middle. Movement further to the left of the diagonal represents an increase in adaptive benefit. The generic case (Figure 4) does not exhibit the pattern in the upper-right, staying mainly parallel to the diagonal. These patterns indicate that adaptive benefit is greater for the extremes (either positive or negative) of adaptive gain values in the individualized case and for the lower adaptive gain values in the generic case.

The histograms in the individualized case tend to shift to the left as the number of bands is increased, while they tend to stay constant in the generic case. Except for some broadening of the histograms in the individualized case, no other significant changes occur with increase in number of bands. This indicates an increasing adaptive benefit with number of bands only for the individualized case.

There are very few moments of negative adaptive benefit. These moments are shown by data to the right of the rightmost diagonal line (note again, the logarithmic scale). More such moments occur in the generic case than in the individualized case, and they tend to decrease in number in the latter case when the number of bands increases. Such moments are due to the finite adaptation rate of the adaptive implementations.

Figure 5 shows box plots of gain for each case. Each panel shows results for a different environment. These are determined assuming a 'smart' hearing aid, that is, one that would switch to omnidirectional in instances where the directional gain would be positive. In such instances the gain is set to 0 dB. Trends in the gain are as expected from examination of the histograms. In particular, except for the individualized adaptive case, there are not large differences in gain statistics as processing type or number of bands changes. Median gain is quite similar in transport and exterior environments, with median values for the interior environment about 1 dB higher than in the other two. The individualized adaptive implementation shows best performance in each condition, although significantly so only in the 256-band case. The 4–7 dB range of median gains in Figure 5 is not much wider than the range expected for two-microphone designs in a diffuse field[4].

Figure 6 shows box plots of adaptive benefit for each case. In the three environments (shaded boxes) in the generic case no 95th percentile of benefit exceeds 1.5 dB; as a percentage of total recording time (unshaded boxes), no generic 95th percentile of benefit exceeds 1.2 dB. This is significant, since all currently available hearing aid directional systems are of the generic type. The 16- and 32-band individualized cases show some instances of benefit greater than 2.0 dB. In total, however, greater than 95% of 16-band individualized benefit is below 2 dB, and 95% of 32-band individualized benefit is less than 2.3 dB. The 256-band individualized case shows the majority (62% in total) of its benefit exceeding 2 dB.

Discussion

Cumulative performance

We believe the patterns of results described above reflect an acoustic input comprising a diffuse or near-diffuse field punctuated by brief periods of directional energy or uncorrelated noise (e.g. wind or microphone noise) in the two channels. This is supported by listening to the recordings, which often have many similar, simultaneously-active independent sources, apparent reverberation, distributed sources (such as large transport vehicles), combinations of these, or quiet periods. The connection between these acoustic characteristics and the data patterns is described in the following paragraphs.

As noted above, the range of median gains includes that expected for diffuse fields. Variance outside this range can be ascribed to directional or uncorrelated energy in the channels. For instance, when a directional source is present at a level high enough to control the adaptive process and does not fall in the null of the static algorithm, the adaptive algorithm can steer a null towards the jammer while the static algorithm cannot. For such occurrences the adaptive algorithm will tend to have lower gains than the static algorithm, causing data to accumulate further to the left in the histograms for lower adaptive gain values. Also, when uncorrelated noises dominate the front and rear channels, the individualized adaptive algorithm will delay the front channel to time-align target-direction energy across the

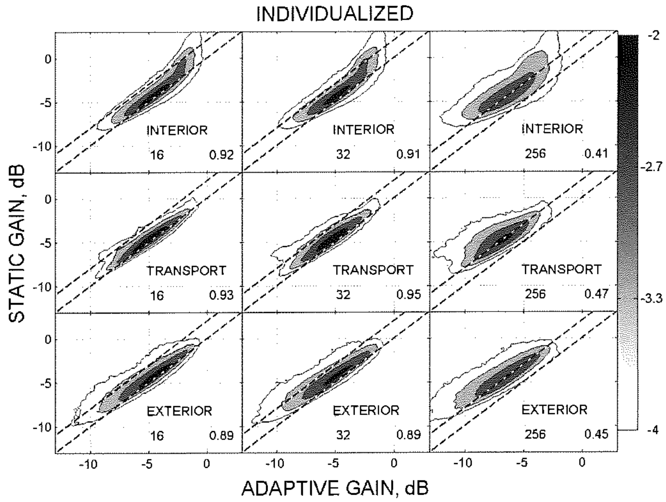

Figure 3. Contour plot of the histogram showing proportion of processing frames attaining certain adaptive/static gain pairs for the individualized processing case. The proportion values on the contours themselves are 10^{-4} (outermost), $10^{-3.3}$, $10^{-2.7}$, and 10^{-2} (innermost; exponents shown by colorbar). Areas of a given shade show gain-pair values that occur at proportions between the values of their surrounding contours. For example, the white area between the outermost and next-outermost contours shows gain pairs that occur with proportions between 10^{-4} and $10^{-3.3}$. Due to the concentrated nature of the data in the histograms the shading corresponds logarithmically to proportion. Each panel shows a different band-number/environment condition. Number of bands is given by the lower, left text value in each panel, and the lower, right text value is the total proportion of frames with gain pairs between the two diagonal lines. The left and right diagonal lines show all gain pairs with an adaptive benefit of 2 dB and 0 dB, respectively.

channels and then average. This yields a −3 dB gain value for uncorrelated signals. The generic and individualized static algorithms implemented here cannot perform this operation because they rely on fixed delays and subtraction followed by amplification (to correct for low-frequency target roll-off caused by the subtraction). Thus, these algorithms can actually *increase* noise levels at their output (though, as noted above, 'smart' hearing aids will turn off directional processing in such cases). The upper-right portion of the histograms for the interior condition is consistent with uncorrelated signals in the two channels.

Other aspects of the patterns are also consistent with combinations of the acoustic field described and processing type. For instance, the consistently small benefit across conditions in the generic case is likely due to similar adaptive and static performance for the diffuse field and uncorrelated noise, and the adaptive implementation's limited ability to take advantage of increased frequency resolution. The latter effect is due to the generic adaptive implementation's directional weights being constrained to a linear combination of two fixed values. Examination of the weights of the individualized adaptive implementation (not shown) for a diffuse input indicates a wide range of weight values that the generic implementation cannot attain.

Moment-by-moment performance: The 'wow' effect

The benefit analysis indicates that relatively few moments of significant benefit (i.e. greater than the 2 dB threshold suggested by Killion (2004)) occur for any but the individualized, 256-band system. Such moments, however, may yield a 'wow' effect, that is, a subjective impression of difference in implementation outcomes much greater than expected from their percentage of occurrence. To help determine from the data whether or not such an effect might exist, Figure 7 shows the number of occurrences of benefit 'streaks' of various durations and threshold values accumulated over the total recording duration (213 minutes). A streak is a time segment during which the benefit continuously exceeds a given threshold value. The streak occurrences are accumulated in bins spaced by a factor of $\sqrt{2}$ and on the horizontal axis, starting at 1 s. For example, the symbol above 4 s on the horizontal axis shows the total number of occurrences of streaks with durations between $4/2^{1/4}$ and $4 \cdot 2^{1/4}$. Unlike the histograms in Figure 6, this analysis allows us to determine whether or not extended moments of significant benefit occur.

If we assume a 'wow' effect is equivalent to a background noise level being reduced by one loudness category (e.g. from 'loud' to 'comfortable but slightly loud') then Figures 2 and 6 of Ricketts and Bentler (1996) indicate that this requires a noise level decrease of approximately 7.5 and 6.2 dB for normal-hearing and moderate-severely hearing impaired listeners, respectively. Figure 7 shows that streaks of such benefit occur only for the 256-band individualized implementation (diamonds and circles, upper right panel), and that they are rare[5] and of short duration. By our definition, then, it is unlikely that a 'wow' effect would be generated by any of the implementations tested. Of the streaks occurring for our generic implementations, only those with 1 dB threshold occur in significant numbers, and most of these are less than six seconds in duration. There are also rare and brief 2-dB threshold streaks for generic implementations, consistent with a lack of perceptible difference in the existing field trials.

Although from this analysis the presence of a 'wow' effect seems unlikely, we have not demonstrated its absence completely since we may not have defined it correctly. While it might be possible to devise some experiment that would help clarify this definition, it is perhaps more instructive to view our data in the context of current field reac-

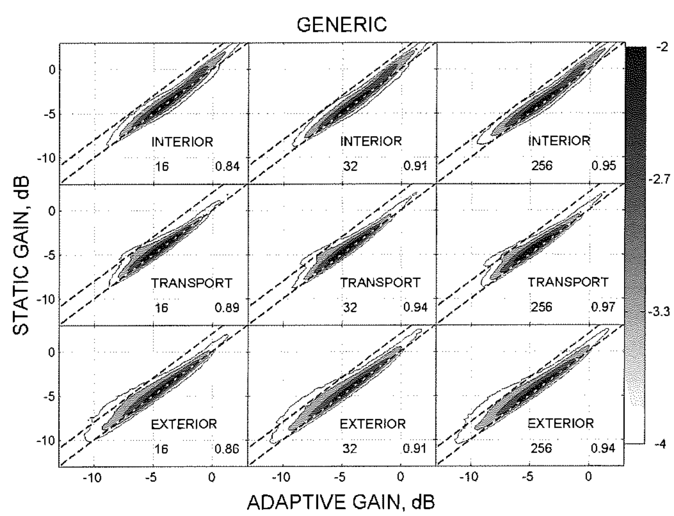

Figure 4. Same as Figure 3, but for generic processing.

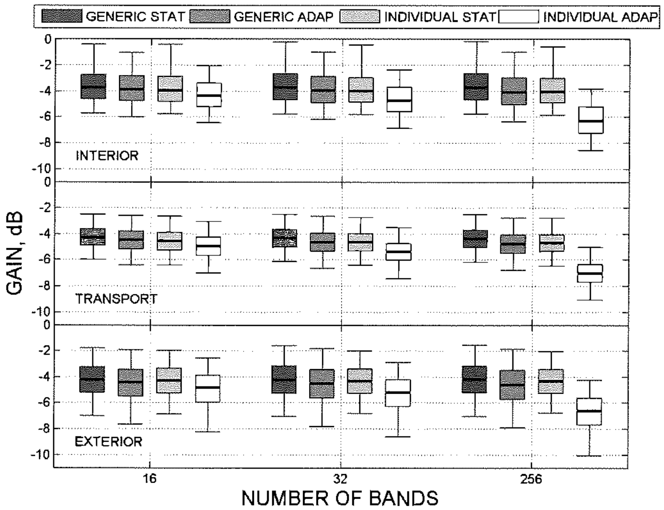

Figure 5. Gain box plots for each condition as a function of number of bands. Gain statistics are determined from gain signals with positive gain values replaced by 0 dB. This simulates the action of a 'smart' hearing aid (see text). Different panels show results from different environments. Results for different implementations are shown by the differently-shaded boxes (see legend). The line inside a box is at the median gain, the lower and upper box borders are at the 1st and 3rd quartile, respectively, and the lower and upper error bars are at the 5th and 95th percentile, respectively.

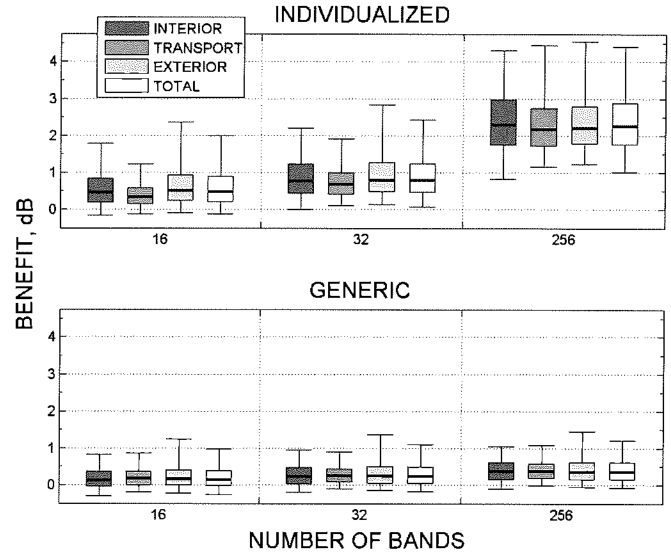

Figure 6. Benefit box plots for each condition as a function of number of bands. Top and bottom rows show results for individualized and generic processing, respectively. Results for different environments are shown by the differently-shaded boxes (see legend). The line inside a box is at the median benefit, the lower and upper box borders are at the 1st and 3rd quartile, respectively, and the lower and upper error bars are at the 5th and 95th percentile, respectively.

tions (Palmer et al, 2006; Bentler et al, 2004) to adaptive directional implementations. Assuming the devices were working appropriately in the field trials and that any 'wow' effect would have generated different outcomes in those studies, it is parsimonious to assume that there would be no such effect in our study as well. This implies that such an effect requires streaks of longer duration, greater benefit, or both longer duration and greater benefit, than the short, 1-dB threshold streaks we find in our data for generic implementations.

Real-world processing constraints

In order to assess the best possible conditions for adaptive benefit, our methods were specifically designed to avoid processing constraints that may have yielded lower benefit. Our examination of individualized arrays is one example of this. In addition, we did not compromise performance by constraining delay through the processing, using speech detectors to stop adaptation, or use only the very short filters currently available. We eliminated channel mismatch because such mismatch can potentially decrease the benefit of two-microphone adaptive arrays as discussed in the Introduction. For these reasons we contend our methods yielded the best possible adaptive benefit obtainable, given our recordings. Still, with these conditions only the highest-resolution, individualized system showed the bulk of benefit to be above Killion's (2004) criterion of 2 dB for perceptible benefit. Practically all (> 95%) benefit for generic systems is below 1.2 dB, and 96% and 92% of benefit is below 2 dB for the 16- and 32-band individualized systems, respectively. Benefit exceeding 2 dB in these latter two cases is brief and rare. Thus the only system obtaining perceptibly significant benefit for a substantial portion of time used filter lengths and a fitting procedure not currently available on the hearing aid market. Realization of such a system in a marketable aid would need to decrease the amount of delay relative to that used here (cf. Stone & Moore, 1999, 2002, 2003, 2005).

Implications for other algorithms

Because of the fidelity of the recordings and calibration and processing in floating point, we maintain that, at the frequency resolutions and adaptation rate tested, the range of benefit found is limited by the environment and not hardware or processing compromises. This implies that our results apply to other similar-resolution/ adaptation-rate two-microphone algorithms that we did not test, including other adaptive directional systems (e.g. the multichannel Wiener filter, Doclo et al, 2005, and systems with microphones at the two ears instead of two microphones at one ear), independent component analysis (e.g. Hyvarinen et al, 2001), and blind source separation algorithms (e.g. Buchner et al, 2005). That is, it can be expected that benefit with these other algorithms would likely not exceed that found here.

The use of faster adaptation rates and/or higher-resolution processing may yield greater benefit than found here, but the potential improvement would need to be traded off against possible concomitant negative effects. For instance, fast adaptation may be deleterious to audio quality or speech intelligibility when combined with head movements. Also, increasing frequency resolution leads to increasing time delay through the hearing aid, which exacerbates problems with unstable feedback and own-voice quality. The effects of design/application mismatch on these systems have also not been investigated with comprehensive real-world data. Adding still more processing may mitigate these effects (e.g. using a voice

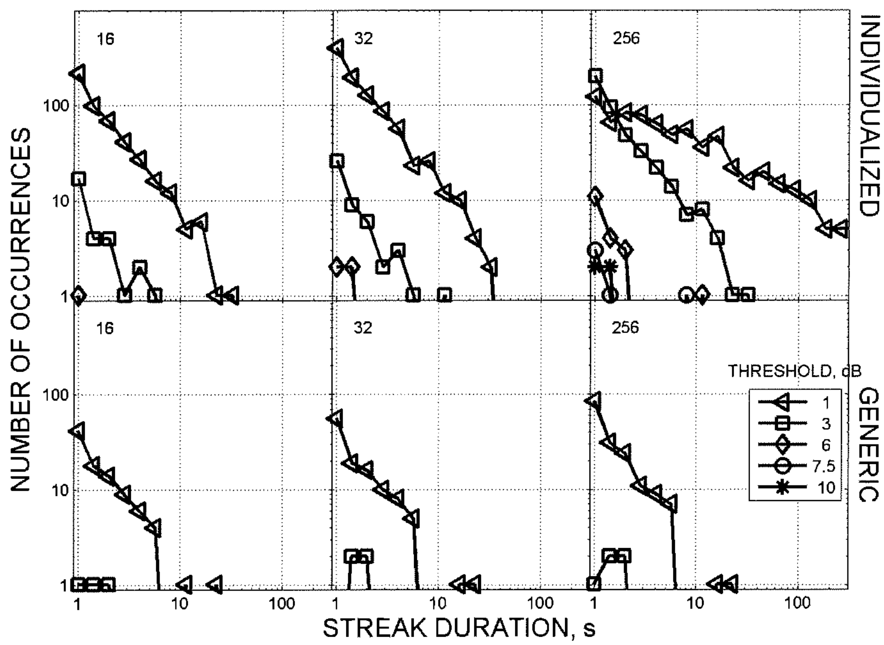

Figure 7. Number of occurrences of streaks of various thresholds as a function of streak duration totaled across all recordings. A streak is a time segment during which the benefit continuously exceeds a given threshold value. The streak occurrences are accumulated in bins spaced by a factor of $\sqrt{2}$ on the horizontal axis, starting at 1 s. For example, the symbol above 4 s on the horizontal axis shows the total number of occurrences of streaks with durations between $4/2^{1/4}$ and $4 \cdot 2^{1/4}$. Top and bottom rows show results for individualized and generic processing, respectively. Left, middle, and right columns show results for 16-, 32-, and 256-band processing, respectively. Legend shows streak threshold in dB for each symbol. Total recording time was 213 minutes.

activity detector to halt adaptation during speech), but doing so may compromise benefit and entails its own trade off against other processing requirements due to the limited capability of today's hearing aids.

Limitations of the study

Our methods imply several limitations of the study. We used recordings from five subjects, each obtained over a one-day period, which limits the range of scenarios encountered in our work. Also, because we obtained recordings with no target signals present, it is not possible to determine benefit for target-alone situations. For example, computing target gain using anechoic targets precludes determining the effect of target reverberation. For these reasons we relied on combining our results with the results from field trials already published (Bentler et al, 2004; Palmer et al, 2006) when discussing general prospects for adaptive technology. In particular, given the greater extent of the field trials and since they did not find perceptible difference between static and adaptive directional systems, we infer that either the benefit is low for the target-alone case or it occurs infrequently enough that it does not increase overall perceived benefit greatly, or both. Of course, these inferences are tempered if the directional processing of the aids in the field studies was limited in some way[6].

Summary and Conclusion

Our results are consistent with the results of field studies (Palmer et al, 2006; Bentler et al, 2004), and theoretical work (Woods & Trine, 2004) indicating small benefit of adaptive null-steering over static directional algorithms in the real world. We found low adaptive benefit overall (<1.2 dB for 95% of recording time), and rare, short periods of any significant (> 2.0 dB) adaptive benefit for generic (i.e. similar to currently marketed) implementations in our analyses of real-world recordings. The only system reaching Killion's (2004) 2-dB threshold of perceptible difference for a significant (> 50%) portion of the recording time was a 256-band system individualized to each subject in the study. The benefit of 16- and 32-band individualized systems was less than 2 dB for 96% and 92% of the time, respectively.

Given our use of real-world recordings and the care taken to ensure proper directional performance, and the results of other field trials cited, it is not unreasonable to conclude that the real-world environment represents a significant limitation on the potential benefit of current two-microphone adaptive directional algorithms. That is, it is likely that there are either few moments of high direct-to-reverberant or direct-to-diffuse energy (which are required for significant adaptive benefit) in real world situations, or that they occur in such situations that subjects do not perceive as a significant benefit any change in aid output due to adaptive directional processing. The former occurs in the presence of multiple simultaneously active sound sources, spatially distributed sources, moderate or greater reverberation, or combinations of these factors—all of which can be heard in our recordings. The latter might occur, for example, if subjects in fact want to switch attention between several high-DRR sound sources without the others being attenuated, such as during a conversation with two or more other individuals in a dry environment.

This is not to say that there will never be real-world moments when an adaptive algorithm will have significant advantage over a fixed algorithm when implemented in the real world. Our recordings spanned only 3.6 hours in total (representing the different scenarios encountered by five subjects over one day, each), so we may have missed such moments. Also, the objective performance measures we used may not have captured moments of subjective significance even if they did exist. Taken all together, however, the evidence from this and other research leads us to conclude—similarly to Walden and colleagues for static directional and omnidirectional comparisons—that signal, noise, and acoustic characteristics combine to yield relatively few moments in which significant real-world adaptive directional processing benefit can be perceived with current technology. Also, given the lack of data on the relative effects of microphone and design/application mismatch on current adaptive and static technologies in real-world contexts, the claim that adaptive technology is always as good as or better than static technology should be viewed skeptically. Such mismatch effects gain increased influence in a comprehensive cost/benefit analysis given a relatively small potential adaptive benefit. Future efforts should consider comprehensive investigation of these effects.

Acknowledgments

A preliminary version of this work was presented at the August, 2008 International Hearing Aid Research Conference in Lake Tahoe, California, USA. We thank two anonymous reviewers for comments that helped improve this work.

Notes

1) In our setup the first room reflection to reach an aid would be from the floor, and would arrive approximately 6.9 ms after the direct wave arrival.
2) In general the convergence time of the adaptive systems used here is on the order of this time constant, i.e. 50 ms. For example, an exponential decay with a 50 ms time constant is a reasonable description of the time-dependence (on a linear time scale) of directional gain in dB in response to a step-change in the incidence angle of an anechoic source from 0 to 180°.
3) Individualized and generic implementations will both sometimes yield non-unity target gain due to limited frequency resolution. Also, this shortcoming is worsened in the generic implementations because they are designed in free space but used on a head.
4) Killion et al (1998) report a 4.7-dB articulation-index weighted directivity index (AIDI) on a manikin, and a hypercardioid is known to generate an AIDI of 6 dB in a free-space diffuse field.
5) The maximum possible number of streaks M, of a given duration D s, is given by the formula $M = T/(D+0.0094)$, where T is the total duration (in seconds) of a given continuous recording. The 0.0094 arises from the fact there are 9.4 ms between frames and there must be one frame of below-threshold benefit between each consecutive streak in the case of the maximum possible number. For our set of recordings, M decreases linearly on a log-log scale from M = 12 660 for D = 1 s, to M = 68 for D = 128 s. M drops more quickly at higher D values, with M = 37 and M = 23 for D = 181 s and D = 256 s, respectively.
6) The field-trial reports show no tests of the aid directionality *per se* (e.g. via polar patterns). Ricketts and Henry (2002), however, demonstrated improved speech intelligibility in noise over omnidirectional processing using aids from the same manufacturer as those used in Bentler et al (2004). Also, Bentler et al (2006) demonstrated improved speech intelligibility in noise over omnidirectional processing using the aids from Palmer et al (2006). These aids, however, only allowed automatic switching between omnidirectional and directional modes when a directional mode was tested. Thus, directionality may not have been active in all instances where benefit for the adaptive case might obtain.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Allen J.B. 1977. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Trans Acoust Speech Signal Process*, 25, 235–238.

ANSI S3.5-1997 1997. Methods for Calculation of the Speech Intelligibility Index. American National Standards Institute, New York.

Bentler R.A., Tubbs J.L., Egge J.L., Flamme G.A. & Dittberner A.B. 2004. Evaluation of an adaptive directional algorithm in a DSP hearing aid, *Am J Audiol*, 13, 73–79.

Bentler R., Palmer C. & Mueller H.G. 2006. Evaluation of a second-order directional microphone hearing aid: I. Speech perception outcomes. *J Am Acad Audiol*, 27, 179–189.

Buchner H., Aichner R. & Kellermann W. 2005. A generalization of blind source separation algorithms for convolutive mixtures based on second-order statistics, *IEEE Trans Acoust Speech Signal Processing*, 13, 120–134.

Cord M., Surr R., Walden B. & Dyrlund O. 2004 Relationship between laboratory measures of directional advantage and everyday success with directional microphone hearing aids. *J Am Acad Audiol*, 15, 353–364.

Cord M., Surr R., Walden B. & Olsen L. 2002. Performance of directional microphone hearing aids in everyday life. *J Am Acad Audiol*, 13, 295–307.

Cox H., Zeskind R.M. & Owen M.M. 1987. Robust adaptive beamforming. *IEEE Trans Acoust Speech Signal Processing*, 35, 1365–1376.

Crochiere R.E. & Rabiner L.R. 1983. *Multirate Digital Signal Processing*. Englewood Cliffs, New Jersey: Prentice-Hall.

Doclo S., Spriet, A., Wouters J. & Moonen M. 2005. Speech distortion weighted multi-channel wiener filtering techniques for noise reduction. In: J. Benesty & S. Makino (eds.) *Speech Enhancement*, NY: Springer, 199–228.

Elko G.W. & Pong A-T. Nguyen. 1995. A simple adaptive first-order differential microphone. in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 169–172.

Foster S. 1986. Impulse response measurement using Golay codes. In: *Proc. IEEE Int Conf on Acoust Sp and Sig Proc*, 929–932.

Greenberg J.E. & Zurek P.M. 1992. Evaluation of an adaptive beamforming method for hearing aids. *J Acoust Soc Am*, 91, 1662–1676.

Hyvarinen A, Karhunen J. & Oja E. 2001. *Independent Component Analysis*. New York: John Wiley & Sons.

Hodgson M. and Nosal E-M. 2002. Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms. *J Acoust Soc Am*, 111, 931–939.

Kates J.M. 2008. *Digital Hearing Aids*. San Diego: Plural Publishing, Inc.

Kates J.M. & Weiss M.R. 1996. A comparison of hearing-aid array-processing techniques. *J Acoust Soc Am*, 99, 3138–3148.

Killion M., Schulein R., Christensen L., Fabry D., Revit L. et al. 1998. Real-world performance of an ITE directional microphone. *Hear J*, 51, 1–6.

Killion M. 2004. Myths about hearing in noise and directional microphones. *Hear Rev*, 14, 16, 18–19, 72–73.

Lockwood M.E., Jones D.L., Bilger R.C., Lansing C.R., O'Brien W.D. et al. 2004. Performance of time- and frequency-domain binaural beamformers based on recorded signals from real rooms. *J Acoust Soc Am*, 115, 379–391.

Palmer C., Bentler R., Mueller H.G. 2006. Evaluation of a second-order directional microphone hearing aid: II. Self-report outcomes. *J Am Acad Audiol*, 27, 190–201.

Pavlovic C.V. 1987. Derivation of primary parameters and procedures for use in speech intelligibility predictions. *J Acoust Soc Am*, 82, 413–422.

Rhebergen K.S. & Versfeld N.J. 2005. A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners. *J Acoust Soc Am*, 117, 2181–2192.

Ricketts T.A. & Bentler R.A. 1996. The effect of test signal type and bandwidth on the categorical scaling of loudness. *J Acoust Soc Am*, 99, 2281–2287.

Ricketts T.A. & Henry P. 2002. Evaluation of an adaptive, directional-microphone hearing aid. *Int J Audiol*, 41, 100–112.

Ricketts T.A. & Hornsby B.W.Y. 2003. Distance and reverberation effects on directional benefit. *Ear Hear*, 24, 472–484.

Stone M.A. & Moore B.C.J. 1999. Tolerable hearing aid delays. I: Estimation of limits imposed by the auditory path alone using simulated hearing losses. *Ear Hear*, 20, 182–192.

Stone M.A. & Moore B.C.J. 2002. Tolerable hearing aids delays. II: Estimation of limits imposed during speech production. *Ear Hear*, 23, 325–338.

Stone M.A. & Moore B.C.J. 2003. Tolerable hearing aids delays. III: Effects on speech production and perception of across-frequency variation in delay. *Ear Hear*, 24, 175–183.

Stone M.A. & Moore B.C.J. 2005. Tolerable hearing aids delays. IV: Effects on subjective disturbance during speech production by hearing-impaired subjects. *Ear Hear*, 26, 225–235.

Walden B, Surr R, Cord M. 2003. Real-world performance of directional microphone hearing aids. *Hear J*, 56, 40–47.

Walden B., Surr R., Cord M., 2005. Real-world performance of directional microphone hearing aids. *Seminars in Hearing*, 26, 70–77.

Walden B., Surr R., Cord M. & Dyrlund O. 2004. Predicting hearing aid microphone preference in everyday listening. *J Am Acad Audiol*, 15, 365–396.

Walden B., Surr R., Cord M., Edwards B. & Olson L. 2000. Comparison of benefits provided by different hearing aid technologies. *J Am Acad Audiol*, 11, 540–560.

Woods W.S. & Trine T.D. 2004. Limitations of theoretical benefit from an adaptive directional algorithm in reverberant environments. *Acoustics Research Letters Online*, 5, 153–157.

Yang W. and Bradley J.S. 2009. Effect of room acoustics on the intelligibility of speech in classrooms for young children. *J Acoust Soc Am*, 125, 922–933.