J Am Acad Audiol 7: 230–239 (1996)

Three Experimental Measures of a Digital Beamforming Signal Processing Algorithm

H. Christopher Schweitzer*†
A. Mark Terry*†
Michael A. Grim*

Abstract

Three studies were conducted to evaluate a prototype nonlinear digital beamforming hearing aid. The experiments all used contextually rich speech materials and subjective analysis methods and, on that basis, are grouped together in this report. In experiment 1, the beamforming was done off line and subjects wore their own hearing aids in an attempt to determine whether the processing reduced the interference of noise jammers. In experiments 2 and 3, subjects used portable, head-worn, real-time devices. The results of these experiments showed improvements in perceived ease of speech understanding for these types of subjective measures and lend encouragement to further development of this type of binaural hearing aid. Additional work is also required to develop methods of physically characterizing the nonlinear directional properties of the algorithms.

Key Words: Beamforming, hearing aid

Clinicians and researchers in recent years have struggled with two parallel realizations. First, speech understanding, or at least sound quality, for hearing aid users is often disproportionately diminished in noisy environments (Plomp, 1978; Plomp and Duquesnoy, 1982; Dubno et al, 1984; Hawkins and Yacullo, 1984; Welze-Mueller and Sattler, 1984; Fabry, 1991; Kochin, 1992; Van Tasell, 1993). Second, there is a need to develop and expand measures of hearing aid performance to better reflect real-world conditions and the new processing capabilities in hearing technology. Cox (1993) expressed this need keenly when she wrote, "Although monosyllabic word tests have been widely used to evaluate hearing aids, their basic unsuitability for this task has been recognized for many years.... it is especially important that the new generation of hearing aids be evaluated with speech test materials that are as natural as possible, and include contextual cues" (p. 299).

As to the technical issue, a number of efforts to improve the functional signal-to-noise ratio (SNR) have concentrated on multimicrophone approaches (Soede et al, 1993a, b; Preves, 1994). Among these methods, "beamforming" is a type of multimicrophone approach that attempts to spatially select acoustic regions that contain a desired signal versus undesired jammers. Various types of filtering approaches for the nonselected regions (Peterson, et al, 1987, 1990; Schwander and Levitt, 1987; Hoffman and Buckley, 1990; Kollmeier and Peissig, 1990; Hoffman et al, 1991; Bodden, 1992, 1994; Greenberg and Zurek, 1992; Kollmeier et al, 1992, 1993; Bilsen, et al, 1993; Stadler and Rabinowitz, 1993; Sullivan and Stern, 1993) have been investigated. Conceptually, the intent is to acoustically "illumine" areas in space of greatest interest to the listener. Hoffman et al (1994) provide a comprehensive summary of the performance properties of spatial filters, or beamformers.

While most research with beamforming algorithms have required laboratory instrumentation, a wearable, real-time beamforming device has been reported by Schweitzer (1994), Schweitzer and Terry (1994), and Terry et al (1994a, b). This type of processing uses computationally intensive algorithms to exploit interaural phase and magnitude to enhance signals arriving from a common direction. In practice, signals arriving from other than 0° azimuth are attenuated. Hence, the wearer operates in a "look to listen" mode. This approach has similarities to those of Kollmeier

*Department of Communication Disorders and Speech Sciences, University of Colorado, Boulder, Colorado; †AudioLogic, Inc., Boulder, Colorado

Reprint requests: Christopher Schweitzer, Chief of Audiology, AudioLogic, Inc., 6655 Lookout Rd., Boulder, CO 80301

and his colleagues (Kollmeier and Peissig, 1990; Kollmeier et al, 1992, 1993).

The AudioLogic algorithm is a nonlinear, frequency domain analysis-synthesis method that is used to modify the separate left and right input spectrums so that frequency components that are identified as originating at positions other than at 0° azimuth are attenuated. The localization of the signal is based on phase and magnitude cues, derived from a fast fourier transform (FFT) analysis of signals obtained from two microphones placed at the ears. These cues of phase and intensity are known to be used by the human auditory system for purposes of localization (Mills, 1972). In this digital implementation of a beamformer, the phase differences between the two microphones for frequencies below 1200 Hz are used to estimate the angle of arrival. At higher frequencies, the phase cue becomes ambiguous because the sound wavelength is less than the width of the human head and a number of signals at different spatial positions can give rise to the same phase difference. This is known as spatial aliasing. For these frequencies, when the phase cue becomes ambiguous, the intensity difference between the two microphones, introduced by the effect of head shadow, is used to estimate localization. For spectral components above 1500 Hz, the intensity difference due to the head-shadowing effect varies from a few dB to 15 dB or more (Sivian and White, 1933). In this implementation, phase and intensity differences are primarily used to identify spectral components that are dominated by noise sources located off of the front-back axis. These components are then attenuated. Thus, intensity ratios between the two ears, measured at a given frequency, that are greater than a certain threshold are used to compute the amount of attenuation applied to the given spectral component. For low-frequency signals, the phase difference, between the two ears, is used to estimate delay and, hence, angle of arrival for that spectral component. For a target sound source straight ahead, the delay would be close to zero. For competing sound sources off to one side, the delay would diverge from zero towards a maximum for sounds at 90° azimuth.

The spectral analysis to estimate phase and intensity of a frequency component was carried out via FFTs. The sampling frequency used in these experiments varied from 14,400 to 19,500 Hz, while the input window length to the FFT was 256 samples. The analysis-synthesis scheme used an overlap-add interval of half of the window length to control for frequency aliasing and for the effect of the windowing. The effective frequency resolution for spectral analysis varied from 56 to 76 Hz. The flow diagram for the nonlinear frequency domain beamformer is shown in Figure 1.

The real-time beamformer used in experiments 2 and 3 was configured to pass sounds arriving from a target sound source in the 0° azimuth, or straight-ahead, position. In effect, the beam gain value for any spectral component was determined by table lookup using the parameters of frequency and computed delay and magnitude ratio between the left and right input spectra for that spectral component. During operation, the beam gain could be optionally smoothed with reference to the gain value from the preceding analysis/synthesis cycle. Front-to-back signals were discriminated by directional microphones, so that signals from directly behind were attenuated by roughly 6 dB relative to straight-ahead signals. Suppression of signals off to the side was accomplished digitally by the algorithm outlined above using the signals obtained from the two microphones positioned at the ears in behind-the-ear (BTE) cases. The

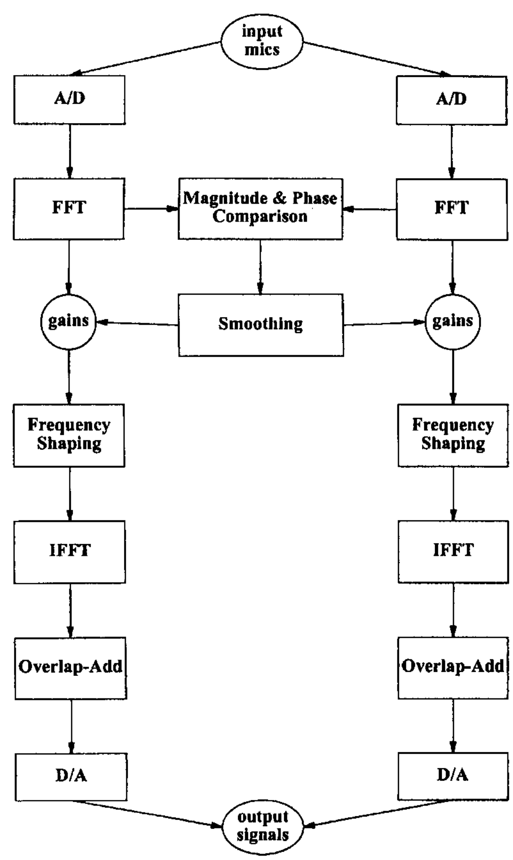

Figure 1 Flow diagram of the beamforming processor used in these experiments.

microphone signals were fed to a box containing a Motorola 56000 DSP board. Following the beamforming, frequency shaping appropriate to the users' hearing loss in each ear was applied in the frequency domain. Optionally, a wide-band compressor could also be applied. The signals were then resynthesized to the time domain. The resulting digital signals were then converted to analog signals and routed to the left and right hearing aid receivers.

This article will report on three studies to evaluate the spatial enhancement processing strategies under development. It should be noted that algorithmic development in this area is continuous, and there is no single beamforming processing algorithm. The basic algorithm implemented was similar in each experiment. However, parameters that controlled beam width and smoothing of the extracted signals were varied. Some of the differences are described later in the discussion section. All of the studies in this series of evaluations made use of contextually rich materials and on that basis were grouped into this report.

EXPERIMENT 1: OFF-LINE SUBJECTIVE MEASURES OF TARGET VOICE DISTINGUISHABILITY

Subjects

In a first attempt to examine the potential usefulness of a beamforming system, 16 hearing-impaired subjects, 11 male and 5 female, were recruited for a subjective listening task using recorded samples of the processing. Subjects were all full-time binaural hearing aid users aged 38 to 80 years with known sensorineural impairments. All subjects had normal tympanometric functions bilaterally. Pure-tone averages for frequencies 0.25 to 4 kHz ranged from 25 to 60 dB HL, and all subjects had generally sloping losses such that frequencies above 1 kHz were more threshold impaired than frequencies below 1 kHz.

Method and Materials

Recordings. Test materials were prepared by mounting two high-quality (Shure Beta 458) microphones on a head-like structure in a large, acoustically unmodified office with moderate (RT60 < 0.5 sec) reverberation. Recordings were then made of four listening scenarios in which "target" talkers competed with various "jammers," a term hereafter used to refer to various types of interference including various competing voices. These recordings of the target voices reading short passages in the presence of the jammers were digitized and processed with and without the beamforming algorithm applied.

The jammers were located at various positions greater than 20° off axis to the front of the artificial head. Rigorous mapping of the jammers was not done for this first attempt to apply beamforming to hearing-impaired listeners. Sound level readings of the taped materials using an Amplaid Model 13 SLM (sound level meter) indicated that the approximate SNR in dBA ranged from +4 dB to +2 dB prior to applying the beamforming algorithm for the segments.

The four scenarios were male talker with single male jammer (M/M), female talker with single male jammer (F/M), female talker with multiple mixed various voice jammers (F/VV), and female talker with loud music and speech (F/LMS). Each segment was approximately 8 seconds long. The tapes were mixed so that the overall levels of each scenario were matched and that the target voice was equally represented in all segments. The task of the postrecording beamforming algorithm was to attempt to extract the target voice by use of the stereo "interaural" acoustic elements.

Subject Task

Play Back Method. The processed materials were presented to the 16 subjects in a standard audiometric sound booth at the University of Colorado's Department of Communication Disorders and Speech Sciences. A Sony digital audiotape (DAT) Model 59ES player was routed through a Grason Stadler 1704 audiometer and transduced by a single Altec Lansing speaker adjusted for an average playback level of 70 dB SPL. Subjects were positioned directly in front of the speaker at a distance of 39 inches.

The subjects wore their own hearing aids throughout the test and were allowed to adjust them to a comfortable level on preliminary material presented for that purpose. After the subjects reported that the basic test level was "in the middle of the comfortable range," no attempt was made to standardize insertion gains or to otherwise control for the various hearing aid differences among the subjects.

Table 1 Experiment 1

	F/M-UB	*F/M-BF*	*F/VV-UB*	*F/VV-BF*	*M/M-UB*	*M/M-BF*	*F/LMS-UB*	*F/LMS-BF*
Question Form 1								
Mean	2.31	1.13	3.13	1.69	3.06	1.19	4.19	2.18
SD	0.87	0.34	1.15	1.08	1.24	0.40	0.83	1.22
	**		**		**		**	
Question Form 2								
Mean	3.00	4.50	1.56	3.69	2.38	4.25	2.75	4.69
SD	1.03	0.73	1.03	0.73	1.15	0.77	1.06	0.48
	**		**		**		**	

**t significant at $p < .001$.

Rating Measures. Two separate 5-point rating judgments were made on the processed materials. The four sound scenarios were presented sequentially with the nonbeamformed, but digitized, segment followed by the beamformed segment of the same passage. Thus, eight ratings were obtained for one rating and eight for another. The first rating asked the question: "How distinguishable is the voice from the noise?" The five alternative rating choices were (1) very distinguishable, (2) moderately distinguishable, (3) somewhat distinguishable, (4) difficult to distinguish, and (5) very difficult to distinguish.

The second rating asked the question: "How much does the noise interfere with the voice?" On this measure, the first rating position was allocated for severe noise interference so that the alternatives were as follows: (1) interferes severely, (2) interferes moderately, (3) interferes somewhat, (4) interferes slightly, and (5) does not interfere at all.

This reversal of the psychometric scaling was done as a means of monitoring intrasubject internal consistency and to determine if the variance, or the main effect, was influenced by the form of the question or by a scaling bias. To control for learning effects, half of the subjects answered question 1 first, and half answered question 2 first. Four practice items consisting of similar, but unscored, segments were given to familiarize the subjects with the task.

Results

Table 1 shows a summary of the mean scores and standard deviations for the 16 sampled items. Recall that for question form 2 (see Table 1) the direction of the ratings was inverted relative to question form 1. Paired t-tests indicated significant differences ($p < .001$) in every case of comparison of the nonbeamed versus beamformed processing for these measures. The mean data for the two question forms are summarized in Figures 2 and 3.

These results indicated that the beamformed material was unambiguously "less difficult" for the group, regardless of how the question was posed. Variance, however, was affected by the sound scenario, as well as by the structure, and/or the scale direction of the five alternative answers. However, regardless of the form of the question, the main effect of beamforming was always significant at the .001 level of confidence.

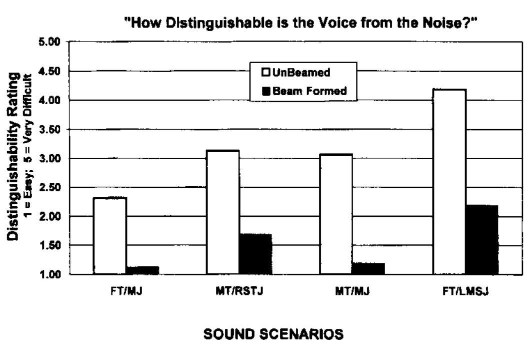

Figure 2 Mean data for question form 1 in experiment 1.

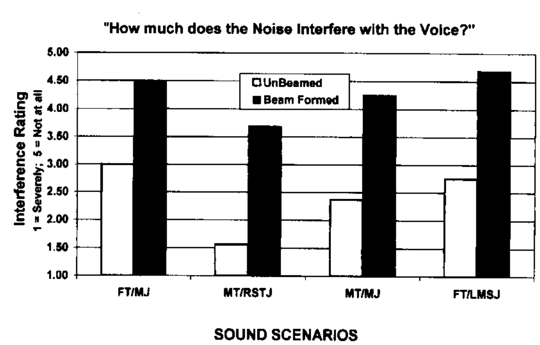

Figure 3 Mean data for question form 2 in experiment 1.

EXPERIMENT 2: REAL-TIME DIFFICULTY RATING (NOISE INTERFERENCE) ESTIMATES

Subjects

Upon development of a real-time version of the beamforming algorithms, another group of 16 hearing-impaired listeners was recruited for studies of the practical effectiveness of the beam. Subject criteria were the same as in experiment 1: bilateral hearing aid users with sensorineural impairments with similar age and hearing loss ranges. Mean age was 53 years. Ten males and six females participated.

Method

Real-time Devices. Portable prototype digital hearing devices were constructed and used for this experiment. Total system delay time for the device was less than 28 msec, and the 14.4 kHz sample rate provided an effective bandwidth of 7 kHz. The wearable devices consisted of BTE-style hearing aid cases in which standard directional (Knowles EL 3077) microphones and receivers (Knowles ED 3088) were mounted. The transducer assemblies were hard-wire connected to the prototype processing unit, which was approximately the size of a "walkman"-style personal cassette player. Soft custom earmolds were constructed for each subject and fit to the BTE elements.

Fitting. By use of an Audioscan RM-500 probe microphone system, the outputs of each subject's hearing aids were matched within ±4 dB throughout the frequency range of 0.25 to 4 kHz. This was done by a software program linking the Audioscan to the Windows-based AudioLogic developmental fitting program. In this study, the volume position of the subjects' own hearing aids was adjusted in order to report that a brief passage of 68 dB SPL of male talker connected discourse was at a level described as "most comfortably clear." This setting was then taken as the gain response that was "matched" by the digital processor.

Materials. This experiment made use of the Subjective Intelligibility Rating (SIR) materials developed by Cox and McDaniel (1989). A digital audiotape (DAT) recording of the 45-second (SIR) speech passages was used as the "target" material. The SIR test uses a male voice reading short topical passages that are essentially equivalent in clarity. The normalized SIR procedure was not conducted in these measures, but the target passages provided a fairly consistent speech source. Jammer noise was provided by four spatially separated voices, two male and two female, recorded on four-track tape. Each jammer voice was recorded reading a passage of text at a fairly rapid pace and professionally mixed for equivalent levels. The jammer voices were then played back on a four-track Tascam TSR-8 1/2-inch reel-to-reel tape system on four separate matched speakers (Polk Audio M42). These were distributed at locations 30 inches to the subject's head at 45°, 135°, 225°, and 335° azimuth.

Subject Task. The intent of the task was to probe again the frequent comment of hearing aid users that it "is more difficult to hear in noise," or that communication in noise requires "a great deal of effort." Hence, this experiment attempted to isolate the subjective "difficulty" aspect of hearing speech in noise. The listener task was to rate each presented passage on a 5-point scale with the following instructions: How difficult is it for you to understand the talker in the center speaker? The five alternative rating choices were (1) not difficult at all, (2) slightly difficult, (3) somewhat difficult, (4) very difficult, and (5) cannot hear the voice. A large, printed version of the task description remained in front of the subjects just below the target speaker during all ratings.

The four talker jammer levels were equalized at the location of the subject's head prior to seating, and the overall level was set and fixed at an average dBA reading of 68. The level of the target voice was varied to create three SNRs of +3, 0, and –3 dB. Upon initial calibration of levels, the target voice level was easily and precisely adjusted by routing the signal from a Sony 59ES DAT player through a Starkey AA30 audiometer set to 1-dB increments of attenuator control. An Infinity RS-125 speaker located 33 inches from the subjects was used for the target voice. Hence, the target speaker was directly forward of the listener and 3 inches more distant than any of the jammer speakers. The level determinations and calibrations were rechecked before and after each session.

A practice period of 30 to 45 seconds, during which the target voice was presented in quiet, was provided to allow subjects to indicate if the basic, unjammed level was too soft or too loud. Slight volume adjustments on either the subjects' own aids (OA) or the experimental device in either beamformed (BF) or unbeamed mode (UB) were observed in a small number of cases.

The order of processing (OA, BF, UB) and the SNR (+3, 0, and –3 dB) were systematically varied between subjects. After the practice session, subjects listened to one of the approximately 45-second passages and indicated their rating to the investigator. Subjects were not deliberately told which of the digital processing schemes were active, but, generally, the attenuation of the off-beam noise is not subtle and, in most cases, it was presumed that there was little doubt whether or not beamforming was activated. Subjects were instructed to look in the direction of the target speaker, but no attempt was made to restrict or control head movements. The starting passage of the SIR material was varied for each subject, but the subsequent order as recorded on the tape was not altered. Hence, the passages per processing condition were not fully randomized, but they were systematically controlled in an attempt to reduce potential differences that might be related to passage content or recording variations. Two passages ("guitar" and "crow") were omitted as the copy of tape used had average levels 2 to 3 dB higher than the remaining passages.

RESULTS

For the group of 16 listeners, mean data were significant in paired t-tests ($p \leq .001$) in each comparison of OA to beamformed processing for the three SNRs. On the other hand, in no case did the experimental device in UB produce a significant difference when compared against the OA condition. The results are summarized in Table 2 and graphically portrayed in Figure 5.

Table 2 Experiment 2: Difficulty Ratings by SNR

	OA	*BF*	*UB*
SNR +3 dB			
Mean DR	3.75	1.66	3.81
SD	0.68	0.67 **	0.68
SNR 0dB			
Mean DR	4.53	2.47	4.41
SD	0.46	0.97 **	0.61
SNR –3 dB			
Mean DR	4.81	3.63	4.69
SD	0.40	0.94 **	0.63

**t significant at p < .001 versus OA.
OA = own aids; BF = beamformed mode; UB = unbeamed mode.

These data suggest that the basic digital processor without beamforming neither increased nor decreased listener perception of difficulty on this task. However, the activation of the beamforming mode clearly did reduce the perceived difficulty in each of the three SNRs; hence, differences from the OA conditions seem reasonably assignable to the operation of the beamformer and not other electroacoustic parameter differences.

Since it was uncontrolled for, there were obviously considerable differences among the

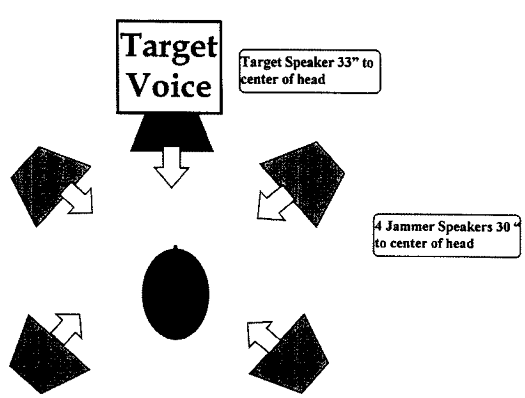

Figure 4 Speaker map for experiment 2 set up. Subjects were located with spatially separated jammer voices in the positions indicated and instructed to rate the forward "target" speaker voice for difficulty.

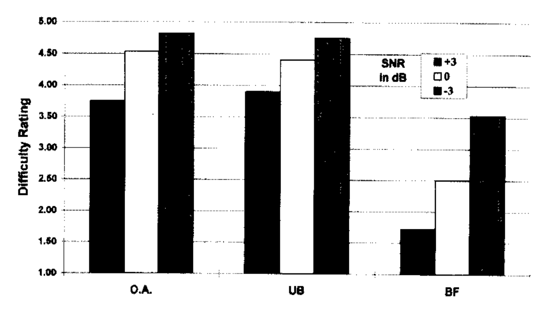

Figure 5 Mean scores for experiment 2. Difficulty ratings for the OA condition and two digitally processed conditions (OA = Own Aids; UB = Unbeamformed; BF = Beamformed) by SNR.

Table 3 Mean Improvements in Experiment 2: Difficulty Rating for Subjects Grouped by Own Aid Type

	Linear	*Standard Compression*	*Two-band Compression*	*All Compression*
Number of Subjects	6	7	3	10
SNR				
+3 dB	2.17	1.93	2.33	2.13
0 dB	2.17	2.00	2.00	2.00
−3 dB	1.25	1.14	1.67	1.41
Mean for All SNR	1.86	1.69	2.00	1.85

16 subjects relative to the types of hearing aids in the OA condition. Six of the subjects had linear hearing aids and 10 wore some version of compression aids. Of those wearing compression aids, three wore programmable devices that were of two-band design. Given the three different styles of amplification among the subjects, it was of ancillary interest whether the "benefit" of the beamforming was impacted by the OA processing condition. In other words, was there a tendency to show less improvement in difficulty rating for those subjects who wore more sophisticated compression aids? Although the small numbers in each group make it difficult to draw any firm conclusion, no systematic tendencies were observed that might have suggested that the base interference from noise in the OA condition was greater for linear hearing aids than for compression. When the relative improvement for type of OA was separated by SNR, no group varied in mean improvement by more than 0.4 rating point. Further, when the improvements were averaged by OA group for all three SNRs, and both types of compression aids were combined into a single group, the average improvements were nearly identical to the linear-aided group (1.86 rating points mean improvement for linear vs 1.85 for all compression-aided subjects). This can be seen in Table 3, which separates the groups on the basis of OA types and shows the mean improvement for the beamformed condition.

EXPERIMENT 3: REAL-TIME SUBJECTIVE INTELLIGIBILITY RATINGS (SIR) MEASURE

The third study in this series was conducted entirely at the University of Colorado as part of a dissertation project (MG).

Subjects

Again, 16 sensorineural subjects with audiologic profiles similar to experiments 1 and 2 participated.

Method

All processing in this study was through the prototype digital hearing device in a linear unbeamformed and a linear (i.e., without compression) beamformed mode.

Fitting. As in experiment 2, the hearing aid "match" was on the basis of a 65 dB SPL input sweep. The basic gain pattern for each fitting was matched within ± 4 dB for the frequency range of 0.25 to 4 kHz using an Audioscan RM-500 probe microphone measurement device coupled to a 386 computer. The real-ear gains of the subjects' hearing aids were measured at volume levels adjusted by the subjects to a comfortable level when listening to a passage of connected discourse presented at 65 dB SPL from a distance of 1 meter prior to the probe measurement sweep. Amplitude dynamic differences (e.g., nonlinear characteristics) were not accounted for in the experimental fittings.

For this study, only one programmable digital device was used on all subjects. This device was an earlier version of hardware than the ones used in experiment 2. It operated at 19.5 kHz sample rate with a total system delay of less than 20 msec.

Test Arrangement. A standard audiometric sound suite with dimensions of 5½ feet by 7½ feet by 6½ feet was used. In order to control for head movements during the lengthy test period, an artificial head was used to mount the BTE microphone pieces. The digital outputs from the processor box were delivered via patch cords to button-style receivers fastened to custom earmolds worn by the subjects. Subjects were seated comfortably a few feet away in the test room, without concern for head or body movements, during this test session. The artificial head was located 1 meter, at 0° azimuth, from the target speaker (Auratone S.S. Cube) used to present the SIR materials. The jammer speaker (Realistic Nova 14) was located at an azimuth of 340° (20° left of center). The elevations for both the target and jammer speakers were at 0° to the center of the artificial head.

Materials and Task. SIR measures were done at SNR of –4 dB and 0 dB for the two processing conditions. The SIR materials were presented via a Nakamichi audiotape player at a level of 65 dB SPL as measured at the artificial head with the head absent. The competing jammer noise was the multitalker babble developed by Cox and McDaniel (1989) for use with the SIR. This was dubbed onto a 12-minute loop and introduced to the jammer speaker by use of a Nakamichi BX-15 cassette deck. The jammer level was adjusted to 61 dB SPL at the center of the artificial head as measured with the head absent from the test point.

Subjects were fully instructed verbally, and by means of written instructions, as to the rating scale procedures used with the SIR test. Each subject scored the test passage by circling the percentage of words understood on a horizontal scale given before each presentation as in Cox and McDaniel (1989). Two practice passages were given in the absence of jammer noise to familiarize subjects with the task. All subjects had 90 percent to 100 percent ratings for the unjammed practice condition. A second practice session used a backwards passage for which all subjects gave a 0 percent rating. Five full-length SIR passages were used for each processing condition and ratings were obtained for each of the processing conditions. The average for the five ratings per condition was used to express the rating score for each condition.

Table 4: Experiment 3: SIR Ratings by Process and SNR

Subject	*Linear –4*	*Beam –4*	*Linear 0*	*Beam 0*
1	78	84	90	94
2	36	63	89	90
3	95	85	99	99
4	66	80	93	90
5	29	38	75	58
6	29	43	65	86
7	36	67	86	93
8	39	53	85	93
9	8	5	47	80
10	2	12	61	54
11	6	6	60	51
12	10	8	49	85
13	10	20	74	42
14	11	26	22	59
15	2	27	82	94
16	10	33	69	88
Mean	29.19	40.63*	71.63	78.50
SD	28.47	28.28	20.39	18.74

*t significant $p < .05$.

RESULTS

As in the previous two experiments, a measurable advantage was observed for the beamformer in these conditions. Table 4 shows the individual performances for the 16 subjects (average of five ratings) for the two SNRs and the two processing conditions. It is obvious from the data that the range of ratings was quite high among the 16 subjects. Standard deviations, as indicated in Table 4, showed that intersubject variance was rather large. The t-test statistics indicated a significant difference ($p < .01$) between the two processing conditions for the –4 dB SNR, but not for the 0 dB SNR. There is a strong possibility of ceiling effects for this condition, which was substantially less difficult than the –4 dB SNR. Average ratings of better than 90 percent can be seen for several subjects in Table 4, and many 100 percent ratings were given in the raw scores prior to averaging.

DISCUSSION

Three separate experiments using rating measures were conducted to evaluate whether digital hearing aid beamforming algorithms can be shown to yield favorable results when applied to hearing-impaired listeners listening to connected discourse in noise. Generally, the SNRs for these tests were rather poor (–4 to +3 dB). This, combined with the absence of normal visual cues, provided for a fairly difficult task of hearing speech in noise for these sensorineural-impaired listeners.

Test conditions and methodologies differed among the three experiments described here. It should be understood that, in each of these three experiments, a different algorithmic version of beamforming software was implemented. Consequently, not only did some properties of the spatial "noise management" processing differ among the experiments, but also other aspects of the electroacoustics were in different stages of development. In experiment 1, the processing was not done in real time, and there was no smoothing of the beam gain between analysis-synthesis cycles. In experiment 2, the processing was done on a real-time device, and smoothing of the beam gain between analysis-synthesis cycles was implemented. Perceptually, the effect of smoothing is most noticed when listening to off-axis sources when the on-line signal is absent. When smoothing is not present, a "waterfall-chirbling" artifact is often heard by normal-hearing listening and sometimes it was reported to be audible to the hearing-impaired subjects. If both target signal and noise "jammer" are coincidentally present, the "chirbling" artifact is reduced and becomes imperceptible. Also, in experiment 2, a wide-band compressor was applied after beamforming together with the appropriate frequency shaping. However, it should be noted that the compressor was also part of the nonbeam processing, and the data show that nonbeamforming was not significantly different from the subjects' OA condition.

In experiment 3, the frequency shaping component was linear, but beam-gain smoothing was applied. In this experiment, the spatial separation between target signal and the single noise jammer was only 20°, which was smaller than in the other experiments. Also, in experiment 3, the multitalker noise jammer was emitted from a single speaker, whereas experiment 2 made use of spatially unique voice jammers; experiment 1 was different than either of the other experiments in that regard. Nevertheless, it seems that our basic investigational question of whether the application of beamforming can be shown to measurably reduce communicative stress (difficulty) or improve subjective ratings of speech understanding for these types of listeners was answered with a "yes" in each of these experiments.

The actual mechanism of the noise reduction process implemented here may involve a number of factors. In essence, the nonlinear frequency domain beamformer operates as a spatial filter. In the current configuration of the beamformer, the target signal is assumed to originate from the straight-ahead 0° azimuth location and competing noise sources from other (off to the side) locations. Using cross-spectral cues, phase difference, and magnitude difference, each spectral component is classified as essentially on-line signal (within the beam) or as noise. In this scheme, separation of signal and noise depends on their respective frequency content, and also on the simultaneous and temporal masking effects of noise components on signal components. In cases where signal and noise components coincide in frequency (i.e., in the current implementation, the same frequency bin in the FFT spectra), attenuation of the noise also implies degradation of the signal. When signal and noise are both speech sources, there will be considerable moment-to-moment variations in SNRs for each spectral component within the speech band. Since speech is a highly redundant signal, considerable degradation of the signal can occur while still maintaining intelligibility, particularly for continuous speech, which also has contextual cues. The overall effect on intelligibility therefore depends on numerous factors, which include a trade-off between reduced masking effect of noise components versus signal degradation. For a hearing-impaired person, there is a possibility of increased susceptibility to masking as well as differing abilities to process a noise-reduced but degraded signal. For these reasons, the operation of the beamformer is not a simple manipulation of SNR and cannot be characterized as a simple attenuation of the noise; neither can the directionality be represented by a conventional polar diagram. The prediction of its effectiveness for a hearing-impaired population calls for detailed modeling. However, the experimental data described here suggest potential benefits from this type of beamforming.

Clearly, investigations into real-world applications must be carried on to further quantify the application benefits and, for that matter, to optimize the implementation of beamforming to individual users. For the studies described here, the beamforming algorithms were applied similarly to all members of the experimental group. It is quite possible that individuals with similar hearing patterns by standard audiometric designations may vary substantially in their requirement and tolerance for spatialized processing schemes. Casual observation during the conduct of these experiments suggests that that may well be the case. The findings were clearly encouraging for continued developmental work in this area and suggest the potential value of hearing aid processing based on binaural models.

REFERENCES

Bilsen F, Soede W, Berkhout A. (1993). Development and assessment of two fixed-array microphones for use with hearing aids. *J Rehabil Res Dev* 30:73–81.

Bodden M. (1992). Cocktail-Party-Processing: concept and results. *Proc Int Congr on Acous*, Beijing.

Bodden M. (1994). *Binaural Hearing and Future Hearing Aids Technology.* Presented at the 3rd French Conference on Acoustics, Toulouse, France.

Cox R. (1993). On the evaluation of a new generation of hearing aids. *J Rehabil Res Dev* 30:297–304.

Cox R, McDaniel D. (1989). Development of the Speech Intelligibility Rating (SIR) test for hearing aid comparisons. *J Speech Hear Res* 32:347–352.

Dubno J, Dirks D, Morgan D. (1984). Effects of age and mild hearing loss on speech recognition in noise. *J Acoust Soc Am* 76:87–96.

Fabry D. (1991). Programmable and automatic noise reduction in existing hearing aids. In: Studebaker G, Bess F, Beck L, eds. *The Vanderbilt Report II.* Parkton, MD: York Press, 65–78.

Greenberg JE, Zurek PM. (1992). Evaluation of an adaptive beamforming method for hearing aids. *J Acoust Soc Am* 91:1662–1676.

Hawkins D, Yacullo W. (1984). Signal-to-noise ratio advantage of binaural hearing aids and directional microphones. *J Speech Hear Disord* 49:409–418.

Hoffman M, Buckley K. (1990). Constrained optimum filtering for multi-microphone digital hearing aids. In: *Proceedings of the 24th Asilomar Conference on Signals, Systems, and Computers.*

Hoffman M, Buckley K, Link M, Soli S. (1991). Robust microphone array processor incorporating head shadow effects. *Proc IEEE Int Conf Acous Speech Sig Proc* ICASSP-91, 3637–3640.

Hoffman M, Trine T, Buckley K, Van Tasell D. (1994). Robust adaptive microphone array processing for hearing aids: realistic speech enhancement. *J Acoust Soc Am* 96:759–770.

Kochin S. (1992). MarkeTrak III identifies key factors in determining consumer satisfaction. *Hear J* 45(8):39–45.

Kollmeier B, Peissig J, Hohmann V. (1992). *Digital Signal Processing for Binaural Hearing Aids.* Presented at the 14th International Conference on Acoustics, Beijing, China.

Kollmeier B, Peissig J, Hohmann V. (1993). Real-time multiband dynamic compression and noise reduction for binaural hearing aids. *J Rehabil Res Dev* 30:82–94.

Kollmeier B, Peissig J. (1990). Speech intelligibility enhancement by interaural magnification. *Acta Otolaryngol Suppl (Stockh)* 469:215–223.

Mills W. (1972). Auditory localization. In: Tobins J, ed. *Foundations of Modern Auditory Theory*, Vol. 2. New York: Academic Press.

Peterson P, Durlach N, Rabinowitz W, Zurek P. (1987). Multi-microphone adaptive beamforming for interference reduction in hearing aids. *J Rehabil Res Dev* 24:103–110.

Peterson P, Wei S, Rabinowitz W, Zurek P. (1990). Robustness of an adaptive beamforming method for hearing aids. *Acta Otolaryngol Suppl (Stockh)* 469:85–90.

Plomp R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J Acoust Soc Am* 63:533–549.

Plomp R, Duquesnoy A. (1982). A model for the Speech-Reception-Threshold in noise without and with a hearing aid. *Scand Audiol Suppl* 15:95–111.

Preves D. (1994). Future trends in hearing aid technology. In Valente M, ed. *Strategies for Selecting and Verifying Hearing Aid Fittings.* New York: Thieme Medical Publishing, 363–393.

Schwander T, Levitt H. (1987). Effect of two-microphone noise reduction on speech recognition by normal-hearing listeners. *J Rehabil Res Dev* 24:87–92.

Schweitzer HC, Terry M. (1994). *Effectiveness of Digital Beamforming Hearing Aid Processing in Real Time and Recorded Simulations.* Presented at the International Congress on Audiology, Halifax, Nova Scotia.

Schweitzer HC. (1994). *Use of Binaural Models in Digital Hearing Aid Applications.* Presented at the International Congress of Audiology, Halifax, Nova Scotia.

Sivian L, White S. (1933). On minimum audible sound fields. *J Acoust Soc Am* 4:288–321.

Soede W, Berhout A, Bilsen F. (1993a). Development of a new hearing instrument based on array technology. *J Acoust Soc Am* 94:785–798.

Soede W, Bilsen F, Berkhout A. (1993b). Assessment of directional microphone array for hearing-impaired listeners. *J Acoust Soc Am* 94:799–808.

Stadler R, Rabinowitz W. (1993). On the potential of fixed arrays for hearing aids. *J Acoust Soc Am* 94:1332–1342.

Sullivan T, Stern R. (1993). Multi-microphone correlation-based processing for robust speech recognition. *ICASSP-93*, II 91–94.

Terry M, Schweitzer HC, Lindemann E, Melanson J. (1994a). Evaluation of a prototype beamforming binaural hearing aid. Presented at the Acoustical Society of America meeting, Cambridge.

Terry M, Schweitzer HC, Shallop J. (1994b). Application of digital beamforming hearing aid to cochlear implant patients. Presented at the American Auditory Society meeting, Halifax.

Van Tasell D. (1993). Hearing loss, speech, and hearing aids. *J Speech Hear Res* 36:228–244.

Welze-Mueller K, Sattler K. (1984). Signal-to-noise threshold with and without hearing aid. *Scand Audiol* 13:283–286.