Performance of an adaptive beamforming noise reduction scheme for hearing aid applications. II. Experimental verification of the predictions

Martin Kompis[a)]
University Clinic of ENT, Head and Neck Surgery, Inselspital, 3010 Bern, Switzerland

Norbert Dillier
ENT-Department, University Hospital of Zurich, 8091 Zurich, Switzerland

(Received 5 November 1999; revised 24 April 2000; revised 25 September 2000; accepted 14 November 2000)

A method to predict the amount of noise reduction which can be achieved using a two-microphone adaptive beamforming noise reduction system for hearing aids [J. Acoust. Soc. Am. **109**, 1123 (2001)] is verified experimentally. 34 experiments are performed in real environments and 58 in simulated environments and the results are compared to the predictions. In all experiments, one noise source and one target signal source are present. Starting from a setting in a moderately reverberant room (reverberation time 0.42 s, volume 34 m^3, distance between listener and either sound source 1 m, length of the adaptive filter 25 ms), eight different parameters of the acoustical environment and three different design parameters of the adaptive beamformer were systematically varied. For those experiments, in which the direct-to-reverberant ratios of the noise signal is +3 dB or less, the difference between the predicted and the measured improvement in signal-to-noise ratio (SNR) is -0.21 ± 0.59 dB for real environments and -0.25 ± 0.51 dB for simulated environments (average±standard deviation). At higher direct-to-reverberant ratios, SNR improvement is systematically underestimated by up to 5.34 dB. The parameters with the greatest influence on the performance of the adaptive beamformer have been found to be the direct-to-reverberant ratio of the noise source, the reverberation time of the acoustic environment, and the length of the adaptive filter. [DOI: 10.1121/1.1338558]

PACS numbers: 43.66.Ts, 43.60.Lq, 43.60.Gk [RVS]

I. INTRODUCTION

Poor speech recognition in noisy environments is a major source of dissatisfaction for numerous users of cochlear implants and conventional hearing aids (Kochkin, 1993; Kiefer *et al.*, 1996). One promising approach to solve this problem is the two-microphone Griffiths–Jim beamformer or adaptive beamformer (Griffiths and Jim, 1982; Peterson *et al.*, 1987), where the signals of two microphones mounted close to the user's ears are postprocessed by an adaptive noise reduction scheme (Widrow *et al.*, 1975). A schematic representation is shown in the lower part of Fig. 1. A detailed description of the adaptive beamformer can be found elsewhere (Peterson *et al.*, 1987; Greenberg and Zurek, 1992; Kompis and Dillier, 2001) and is not repeated here. Numerous experiments have already been performed with this method, showing a wide range of signal-to-noise-ratio (SNR) improvements of 0 to 30 dB (Peterson *et al.*, 1987; Peterson *et al.*, 1990; Greenberg and Zurek, 1992; Dillier *et al.*, 1993; van Hoesel and Clark, 1995; Hamacher *et al.*, 1996). Comparison of these data is difficult because of the different experimental settings and a lack of theoretical background to estimate the contribution of each of these differences on the results. In the companion paper (Kompis and Dillier, 2001), a theoretical framework has been presented, which allows the prediction of the noise reduction that can be expected from an adaptive beamforming noise reduction system in different acoustic settings. However, this framework by itself is of limited value only for two reasons. First, the predictions have not been validated by comparisons to experimental results. Second, as the prediction is a complex function of 11 input parameters, it is still relatively difficult to gain a concept of the complex behavior of the beamformer without systematic variations of all parameters. It is the aim of this investigation to start to close both of these gaps.

II. METHODS

The computer program presented in the companion paper estimates the SNR improvement which can be expected to be reached by an adaptive beamformer as a function of 11 acoustic and design parameters. It is not possible to test all parameter combinations with a reasonable number of different values for each of the 11 parameters. Instead, a different approach was chosen, in which a realistic experimental setting was defined first, from which each parameter is varied separately toward both greater and smaller values. In this way, the number of experiments was reduced to a manageable 92. Throughout this text, the experimental setting from which all parameters are varied will be called central setting.

A. Central setting

The definition of the central setting includes the room, the noise, and the signal source, and a set of design parameters of the adaptive beamformer. Figure 1 shows a sche-

[a)]Electronic mail: martin.kompis@insel.ch

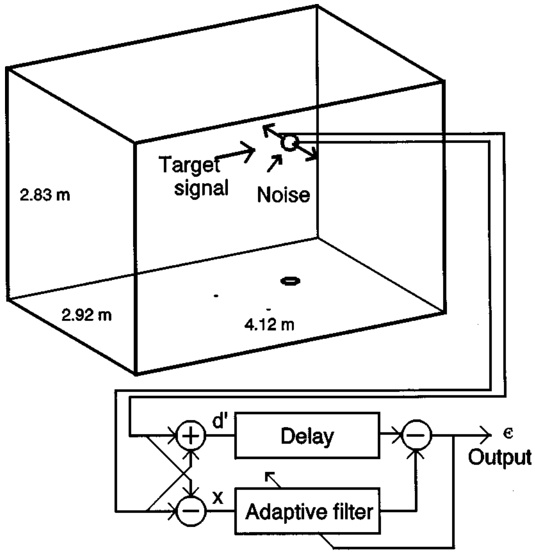

FIG. 1. Schematic drawing of the experimental setup at the central setting (top) and the adaptive beamformer (bottom). The starting points of the arrows in the upper portion of the diagram denote the locations of the loudspeakers and microphones.

matic representation. The parameters of the central setting have been chosen to represent a realistic situation, in which variations of all relevant parameters toward both higher and lower values appear to be reasonable. To define the properties of a suitable room, the dimensions and reverberation times of 18 different rooms (4 offices, 11 living or bedrooms, 1 bath, 2 kitchens) were measured. In this limited sample, the average volume was 34.1 m^3 and the average reverberation time (measured in octave bands with center frequencies of 125–4000 Hz) was 0.41 s and almost frequency independent. These average values may differ, e.g., in a different cultural context. One of these 18 rooms, a shoebox-shaped room with a volume of 34.0 m^3 and an almost frequency independent reverberation time of 0.42 s, was available for experiments for a limited time and was used for the central setting.

Two loudspeakers (Phillips 22AHS86/16R) were placed at a distance of 1 m from a dummy head equipped with a stereo microphone, both from a Sennheiser MKE 2002 set. The azimuth of the loudspeaker emitting the target signal was 0° (i.e., in front of the dummy head), the noise source was 45° to its right. The index of directionality of the loudspeakers was estimated to be 3.4 for band-limited noise 125 and 5000 Hz in an earlier work (Kompis and Dillier, 1993). As to the adaptive beamformer, a sampling rate of 10 240 Hz, a filter length of 25 ms (256 filter coefficients), and a delay of 50% of the filter length (12.5 ms) in the target signal path (marked d' in Fig. 1) was chosen.

B. Experiments in real and simulated rooms

As far as possible, experiments were performed in the real room described. Some of the experimental parameters, most notably the volume of the room and the reverberation time, cannot be readily varied independently using real rooms. Furthermore, the room which was used for the central setting was available only for a limited time for recordings. For these reasons, 58 of the 92 experiments were performed in simulated rooms, using a simulation method presented earlier (Kompis and Dillier, 1993). This simulation procedure is based on an image method introduced by Allen and Berkley (1979). It simulates the impulse responses between acoustic sources and microphones in shoebox-shaped rooms, taking into account the effects of directional sound sources and the acoustic head-shadow of the listener. The head is modeled as a rigid sphere. For the simulations, a value of 18.6 cm was chosen for the diameter of this sphere, as proposed by Kuhn (1977) and used in an earlier study (Kompis and Dillier, 1993). The index of directionality of 3.4 for the two sound sources was approximated by an opening angle of ±65°. Within this opening angle, the signal is emitted equally into all directions, and no signal is emitted outside this angle. For the simulated version of the central setting, all other simulation parameters (i.e., reverberation time, room dimensions, relative positions of the sound sources and the listener) were the same as the corresponding parameters of the real room. The suitability of the simulation method for the purpose at hand was validated in the first experiment (Sec. III A). For the prediction of the performance of the adaptive beamformer, the same set of input parameters was used for both the real and the simulated central setting. Table I shows a synopsis of these input parameters.

TABLE I. Synopsis of the input parameters used to predict the performance for the central setting.

Parameter	Value
Room size V	34 m^3
Reverberation time T_r	0.42 s
Distance listener to target signal source l_s	1 m
Index of directionality of target signal source γ_s	3.4
Alignment factor of target signal source A	4
Distance listener to noise source l_n	1 m
Index of directionality of noise source γ_n	3.4
Azimuth of noise source α_n	45°
Sampling rate F_s	10 240 Hz
Number of coefficients in adaptive filter N	256
Delay in target signal path Δ	128 samples

As the signals from the real and simulated central setting were used for several experiments with different values of the design parameters of the adaptive beamformer, only 14 different sets of recordings in real environments and 37 different sets of simulated signals were used. Table II shows a synopsis of the experiments and environments used.

C. Signal acquisition and signal processing

Target and noise signals were recorded (or for the experiments in simulated rooms: simulated) separately. According to the paradigm used for the prediction of the improvement of the signal-to-noise ratio (SNR) described previously (Kompis and Dillier, 2001), the signals of both the noise and the target signal source were white noise. White noise was generated on a computer and played back via a digital audio tape (DAT) recorder driving only one of the two loudspeakers at a time. Uncorrelated noise sequences of 3 s duration were used for the target and the noise signals, respectively. Recordings of the microphone signals at the dummy head were digitized at a sampling rate of 10 240 Hz

TABLE II. Synopsis of the number of experiments in real and simulated environments.

Parameter varied	Experiments in real environments	Experiments in simulated environments	Sets of recorded signals	Sets of simulated signals
None (central setting)	1	1	1	1
Azimuth of noise signal source (Fig. 3)	5[a]	...	5[b]	...
Distance to noise source (Fig. 4)	2[a]	4	2	4
Index of directionality of noise source (Fig. 5)	1[a]	6	1	6
Alignment factor (Fig. 6)	1[a]	3	1	3
Distance to target signal source (Fig. 7)	2[a]	4	2	4
Index of directionality of target signal source (Fig. 8)	1[a]	6	1	6
Reverberation time (Fig. 9)	1[a]	7[c]	1	7
Room size (Fig. 10)	...	6	...	6
Filter length (Fig. 11)	4[a]	...	0[b]	...
Delay in target signal path (Fig. 12)	16[a,d]	17[c]	0[b]	0[b]
Sampling rate (Fig. 13)	...	4[c]	...	0[b]
Total	34	58	14	37

[a]In addition, results from the real central setting are shown in the corresponding figure.
[b]Recorded or simulated input signals used are identical to those at central setting.
[c]In addition, results from the simulated central setting are shown in the corresponding figure.
[d]In addition, results of one experiment already shown in Fig. 11 (filter length 512, delay 50%) is shown in Fig. 12.

into a computer using a custom-built 12 bit stereo analog-to-digital converter and appropriate antialiasing filters. The spectra of the recorded signals were found to rise slightly toward higher frequencies. Although the effect of this spectral feature on the SNR improvement was found to be small, a two-coefficient finite impulse-response filter (coefficients 0.5 and 0.65) was used to equalize the spectra to within 2 dB in the frequency range of 125–4900 Hz. For experiments in simulated environments, white noise was directly filtered by the impulse responses generated by the room simulation program. The spectra of the microphone signals in the simulations were found to be flat to within 2 dB without further conditioning.

To measure the gain in signal-to-noise ratio, first the recording of the noise signal alone was processed by an adaptive beamforming algorithm. The adaptation of the filter was performed using a normalized least-mean squares algorithm (Bellanger, 1987), where the adaptation time constant was chosen to be 0.2 s. After 2 s the filter was assumed to be in an adapted state and the signal variance in the following second was used as a measure of the variance of the noise signal at the output of the beamformer. The adapted filter was temporarily stored and used in a second run, where the recorded or simulated target signal was processed alone, with the adaptation disabled, i.e., maintaining the adapted coefficients from the first run. Again, the variance of the output signal during 1 s was used as a measure of the variance of the target signal at the output of the adaptive beamformer. Using this procedure, a perfect target-signal detection scheme, which prevents any filter adaptation while a target signal is present, is mimicked. Several such schemes have been proposed (Van Compernolle, 1990; Greenberg and Zurek, 1992; Dillier *et al.*, 1993; Kompis *et al.*, 1997) and used in experiments, and the theoretical analysis and prediction of SNR improvement (Kompis and Dillier, 2001) is based on the assumption that one of these schemes is employed. For all experiments involving longer filters (>25 ms), filter adaptation was allowed for 4 s instead of only 2 s to compensate for the proportionally longer adaptation time constant.

D. SNR improvement and intelligibility-weighted gain

As the investigated prediction method predicts SNR improvement, this measure is used to represent the experimental results. However, for hearing aid applications, the primary goal is improved speech intelligibility and not improved SNR. Because some frequency bands contribute more to speech intelligibility than others, SNR improvement may correlate poorly with improvement in speech recognition, if substantial differences between SNR improvements in different frequency bands do exist. To estimate this effect on the presented data, all experimental results which were compared to the theoretical predictions were also examined by an intelligibility-weighted measure proposed by Greeberg *et al.* (1993). To calculate this intelligibility-weighted gain, signal-to-noise ratios were calculated in 15 one-third-octave bands with center frequencies between 200 and 5000 Hz and weighted according to their contribution to the articulation index (ANSI, 1969).

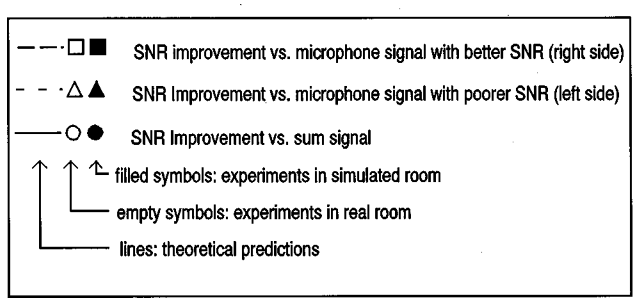

FIG. 2. Legend for the symbols and lines used in Figs. 3–14.

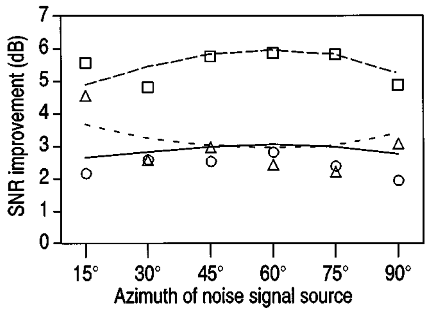

FIG. 3. Influence of the azimuth of the noise signal source. See Fig. 2 for a legend of the symbols used.

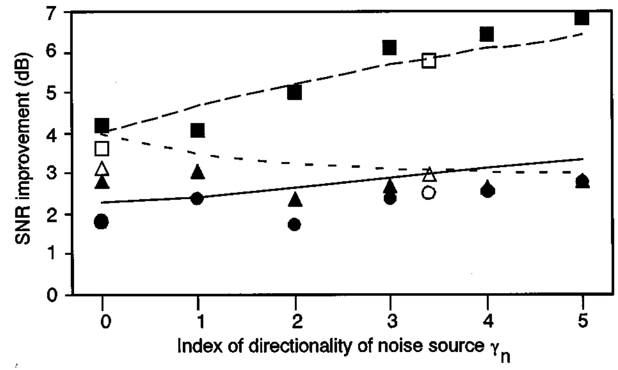

FIG. 5. Influence of the directionality of the noise source. See Fig. 2 for a legend of the symbols used.

E. Representation of the results

The computer program (Kompis and Dillier, 2001) implementing the prediction of the SNR improvement of the adaptive beamformer calculates three different numbers: the SNR improvement versus the microphone signal with the more favorable SNR, the SNR improvement versus the microphone signal with the poorer SNR, and the improvement versus the sum of both microphone signals. The latter corresponds to a simple two-microphone beamformer with fixed postprocessing.

To allow direct comparison, the results of the experiments are similarly calculated as improvements versus each microphone signal and the sum of both microphone signals. Therefore, six sets of data are shown in the figures of Sec. III. All predicted improvements are connected by different lines, and all results from experiments are shown as individual symbols. Results from experiments in real environments are shown using open symbols; results from experiments in simulated environments are shown using closed symbols. Figure 2 shows a legend for all lines and symbols used in Figs. 3–14. SNR improvements at the output of the adaptive beamformer, compared to the SNR at the left microphone (opposite from the noise source and therefore more favorable SNR; triangles in figures) will be lower than SNR improvement versus the right microphone facing the noise source (poorer SNR; squares in figures).

III. RESULTS

A. Results at central setting

At central setting, the predicted SNR improvement was compared to the results from both the experiments in the real room and in the simulated environment. Table III shows the results. All three sets of SNR improvements, i.e., the prediction and the two sets of experimental results, are within 0.5 dB within each other with absolute values ranging from 2.50 to 5.97 dB. None of the data sets exhibit systematically higher or lower values for all SNR improvements when compared to the other two sets.

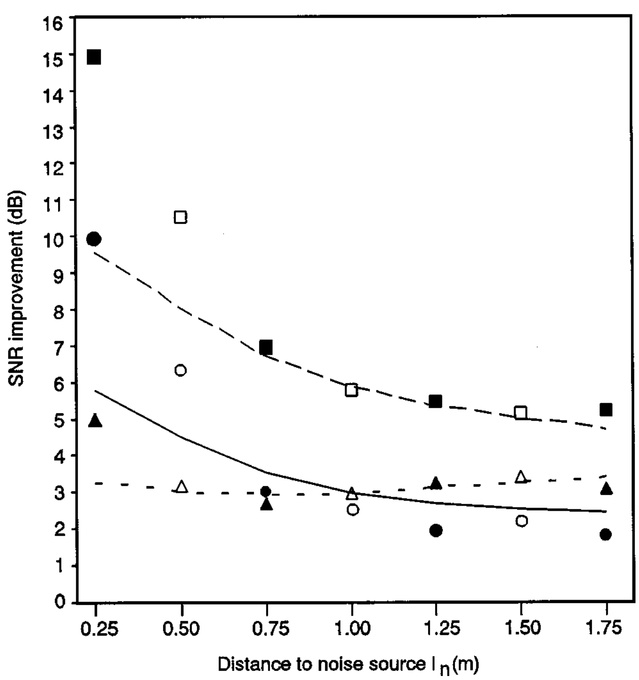

FIG. 4. Influence of the distance between noise source and listener. See Fig. 2 for a legend of the symbols used.

B. Effect of the acoustic parameters of the noise signal source

The three parameters characterizing the noise signal source are its azimuth α_n (with $\alpha_n=0$ defined as the forward direction of the listener), the distance between noise source

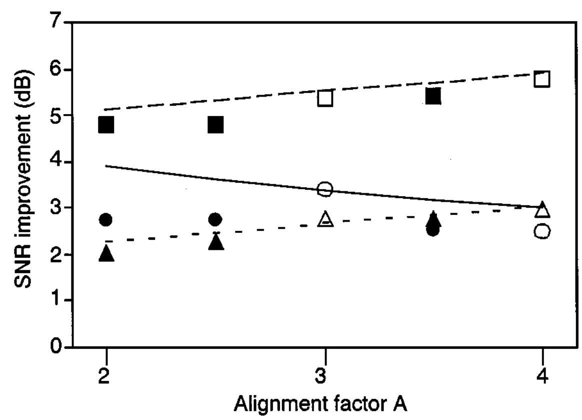

FIG. 6. Influence of the alignment of the target signal source. See the text for the definition of the alignment factor A. See Fig. 2 for a legend of the symbols used.

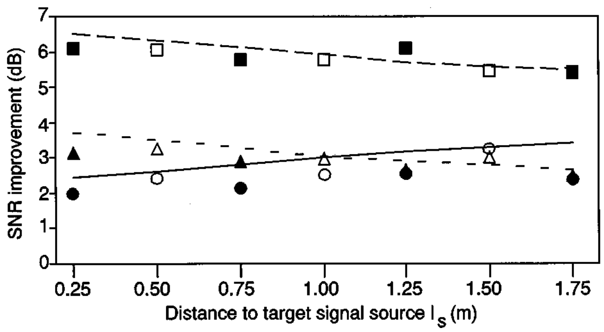

FIG. 7. Influence of the distance between target signal source and listener. See Fig. 2 for a legend of the symbols used.

and listener l_n, and the index of directionality γ_n of the noise source. The last two factors influence the direct-to-reverberant $P_{\mathrm{d/r}}$ ratio of the noise signal. The index of directionality is defined as the ratio between the signal intensity emitted in the direction of the listener to the intensity of a hypothetical omnidirectional source with the same total acoustic output power (DeBrunner and McKinney, 1995).

Figure 3 shows the results of the experiments at 6 different angles of incidence between 15° and 90°. Because of the symmetry of the setting, results can be extrapolated for all angles in the horizontal plane, except for the front (0°) and the rear (180°), where the adaptive beamformer assumes the position of a target- and not of a noise-signal source. The largest difference between prediction and experimental results is 0.89 dB, with more than half of the experimental results lying within 0.5 dB of the predictions.

Figure 4 shows the SNR improvement as a function of the distance between listener and noise source. For distances of 0.75 m ($P_{\mathrm{d/r}}=2.0$ dB) and more, predictions and experimental results are reasonably in accordance. At distances of 0.5 m ($P_{\mathrm{d/r}}=5.5$ dB) and less, however, predictions considerably underestimate the SNR improvement which can actually be achieved using the adaptive beamformer. At $l_n=0.25$ m, the difference is as large as 5.34 dB.

Figure 5 shows the SNR improvement as a function of the index of directionality γ_n of the noise source. For the experiments in the real acoustic environment, $\gamma_n=3.4$ corresponds to the loudspeaker facing the dummy head (central

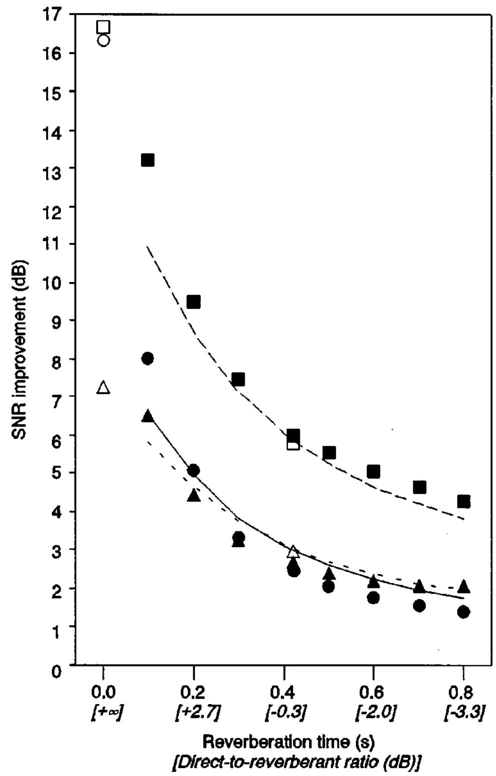

FIG. 9. Influence of the reverberation time. See Fig. 2 for a legend of the symbols used.

setting), whereas $\gamma_n=0$ was approximated by turning the loudspeaker away from the dummy head. For the simulations, $\gamma_n=1$, 2, 3, 4, and 5 was approximated similar to the central setting with opening angles of $\pm 180°$, $\pm 90°$, $\pm 70°$, $\pm 60°$, $\pm 53°$, respectively. For $\gamma_n=0$, an opening angle of $\pm 90°$ facing *away* from the dummy head was used.

There is a reasonable agreement between the predicted SNR improvement and the results of the actual measurements, with an average error of 0.51 dB (range 0.02–1.14 dB). SNR improvement increases with γ_n when compared to the sum signal and to the microphone signal with the less favorable SNR, but decreases slightly when compared to the microphone signal with the higher SNR, presumably due to the increased direct-to-reverberant ratio of the noise signal.

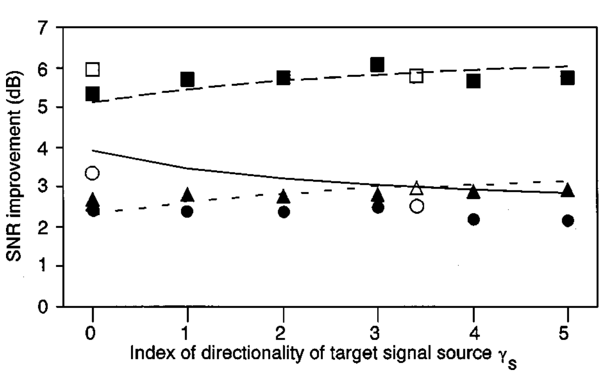

FIG. 8. Influence of the directionality of the target signal source. See Fig. 2 for a legend of the symbols used.

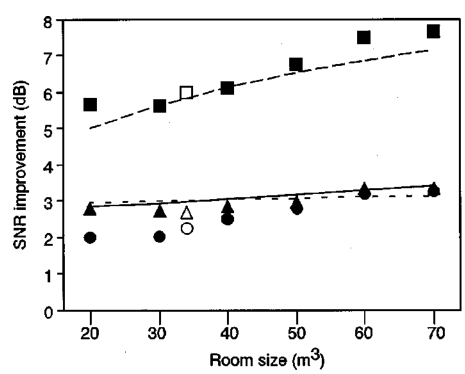

FIG. 10. Influence of the room size. See Fig. 2 for a legend of the symbols used.

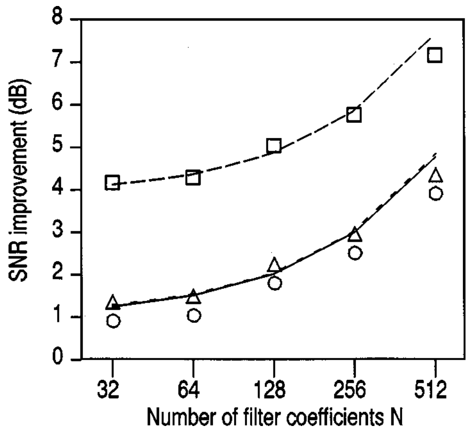

FIG. 11. Influence of the length of the adaptive filter. See Fig. 2 for a legend of the symbols used.

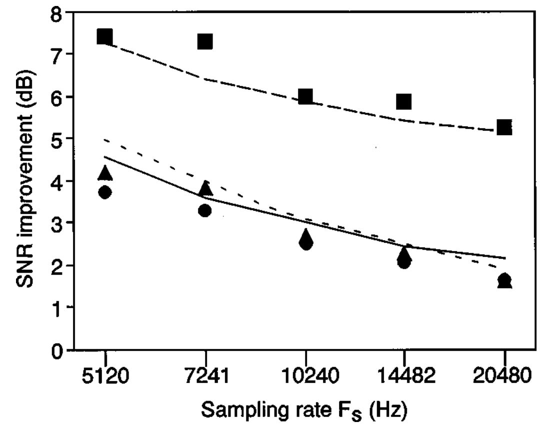

FIG. 13. Influence of the sampling rate. See Fig. 2 for a legend of the symbols used.

C. Effect of the acoustic parameters of the target signal source

The three parameters describing the target signal source are the alignment factor A, the distance between dummy head and target signal source l_s, and the index of directionality of the target signal source γ_s.

The alignment factor A is defined as the ratio between the variance of the nonreverberant portion of a white noise signal after summation of the microphone signals (signal d' in Fig. 1) and the sum of the variances of the two individual microphone signals [cf. Kompis and Dillier (2001) for a detailed discussion]. For perfect alignment, i.e., if there is no delay between the nonreverberant part of the target at the two microphones, A is 4. For a head-sized spacing between microphones and a sampling rate of 10 240 Hz, A drops to 2 (no alignment) for azimuths of approximately 8° and more. The alignment factor can be directly measured in anechoic (real or simulated) environments. For the given setting, the alignment factor was found to be 4 at an azimuth α_s of 0°, 3.5 at $\alpha_s = 3°$, 3 at $\alpha_s = 5°$, 2.5 at $\alpha_s = 6°$, and 2 at $\alpha_s = 8°$.

Figure 6 shows the comparison between predictions and experimental results. The values are in reasonable agreement, i.e., within 0.5 dB when taking either one microphone signal as a reference. For $A = 2$ and $A = 2.5$, the agreement between predicted and measured SNR improvement versus the sum signal differ by 0.88 and 1.15 dB, respectively. The reason for this difference is not completely clear, but most probably a result of the relatively simple model of the direct and reverberant signal parts used to predict the SNR improvement (Kompis and Dillier, 2001).

Figure 7 shows the dependence of the SNR improvement as a function of the distance l_s between listener and the target sound source. Again, the majority of all measured values lie within 0.5 dB of the predictions and display the same tendencies (i.e., SNR improvements decreasing with the distance when taking the microphone signals as a reference, but

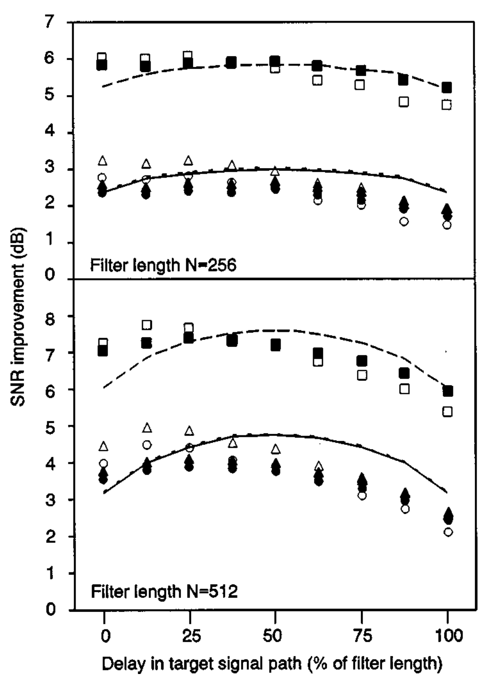

FIG. 12. Influence of the delay in the target signal path. See Fig. 2 for a legend of the symbols used.

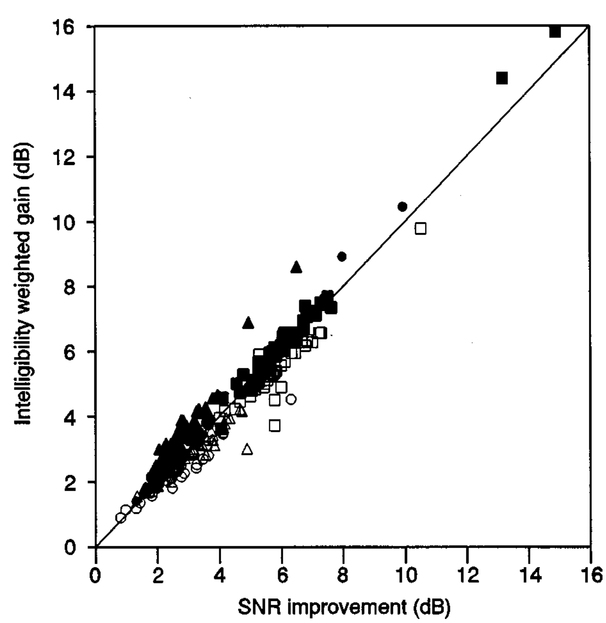

FIG. 14. SNR improvement vs Intelligibility-weighted gain for the experiments in real and simulated environments. See Fig. 2 for a legend of the symbols used.

TABLE III. SNR Improvement at central setting.

	Improvement vs microphone with better SNR (dB)	Improvement vs microphone with poorer SNR (dB)	Improvement vs sum of microphone signals (dB)
Model prediction	3.03	5.85	2.99
Experiment in real room	2.97	5.77	2.52
Experiment in simulated room	2.69	5.97	2.50

increasing when taking the sum signal as a reference). For the given range of distances between 0.25 and 1.75 m, SNR improvements change on the order of magnitude of 1 dB.

Figure 8 shows the dependence of the SNR improvement as a function of the index of directionality of the target signal source γ_s. To obtain the different values for γ_s between 0 and 5, exactly the same procedures were used as for the noise source in Sec. III C. Again, there is a reasonable agreement between prediction and measurement for most data points, with the greatest differences lying toward small γ_s and SNR improvements versus the sum of the microphone signals. For the given range of values $\gamma_s = 0-5$, SNR improvements change on the order of magnitude of 1 dB.

D. Room size and reverberation time

The two acoustical parameters of the room considered in the theoretical analysis of the performance of the adaptive beamformer are volume V of the room and reverberation time T_r. Reverberation time is the time required for a reverberant signal to decay by 60 dB. To vary these two parameters independently, experiments were performed predominately in simulated rooms. Apart from the central setting, only one experiment was performed in a real, anechoic environment.

Figure 9 shows the dependence of the SNR improvement as a function of reverberation time. SNR improvement increases rapidly at short reverberation times. Theoretical predictions and results from the experiments are in reasonable agreement for reverberation times of approximately 0.2 s and above. For shorter reverberation times, the predictions systematically underestimate SNR improvements by up to 2.28 dB at $T_r = 0.1$ s. At $T_r = 0$ s (anechoic environment) the prediction cannot be calculated, as one of the underlying assumptions, the existence of a reverberant signal portion, is violated. The direct-to-reverberant ratio of the noise source is approximately 5.7 dB at $T_r = 0.1$ s, 2.7 dB at $T_r = 0.2$ s, and 1.0 dB at $T_r = 0.3$ s.

Figure 10 shows SNR improvement by the adaptive beamformer as a function of room size. For these experiments, rooms with volumes of 20–70 m^3, in steps of 10 m^3 were simulated. The reverberation time T_r of all these rooms was 0.42 s. To keep T_r constant, the absorption coefficients of the simulated rooms were higher for the larger rooms. Note that for the same absorption coefficients for all rooms, reverberation time would have increased with room size, as everyday experience suggests. As the direct-to-reverberant ratio increases with room size (-2.8 dB at $V = 20\,m^3$, $+2.7$ dB at $V = 70\,m^3$), SNR improvement, especially when compared to the microphone signal with the lower signal-to-noise ratio, increases by approximately 2 dB.

E. Design parameters of the adaptive beamformer

In the theoretical analysis (Kompis and Dillier, 2001), the influence of three design parameters of the adaptive beamformer is considered: the number of filter coefficients in the adaptive filter N, the delay in the sum signal path, and the sampling rate F_s of the system. In practical situations, there will be additional design parameters which influence the performance of a given beamformer, such as the adaptation time constant, the resolution of the analog-to-digital converters, and the performance of any target-signal-detection/adaptation inhibition scheme to prevent filter adaptation in the presence of target signal and therefore target signal cancellation. However, in the theoretical analysis and as a consequence in this study, these additional factors are assumed to be ideal, i.e., filter adaptation is perfect and occurred in the presence of the noise signal only, and all implementation issues are considered to be negligible.

Figure 11 shows SNR improvement as a function of the number N of coefficients in the adaptive filter. In the literature in similar noise reduction algorithms filter lengths of up to 40 ms (Peterson *et al.*, 1987) have been reported. At a sampling rate of 10 240 Hz, this corresponds to a range of $N = 410$ coefficients. In this study, filter lengths between $N = 32(3.125\,\text{ms})$ and $N = 512(50\,\text{ms})$ have been studied. It can be seen that the amount of SNR improvement increases substantially with filter length, especially for values of N of 128 and more. With short filters, e.g., $N = 32$ or $N = 64$, the adaptive beamformer practically routes the microphone signal with the more favorable SNR to the output, but provides only relatively little (approximately 1 dB) SNR improvement above that. In Fig. 11 this is shown by the SNR improvement versus the microphone with the lower SNR improving by more than 4 dB, but only by just above 1 dB when compared to the other microphone signal. For long filters ($N = 512$), the SNR is improved by more than 4 dB, even when compared to the microphone signal with the *more* favorable SNR. The agreement between experimental results and the predictions is reasonable for the entire range of $N = 32-512$, and is poorest for the longest filter.

The theoretical analysis predicts that influence of the delay in the target signal path d' on the amount of noise reduction depends on the length of the adaptive filter N. For this reason, experiments with delays between 0% and 100% of the filter lengths and two different filter lengths ($N = 256$ and $N = 512$ coefficients) were performed for both the real and the simulated central setting. Figure 12 shows the results. It can be seen that for the shorter filter, a variation of only 0.6 dB in SNR improvement is predicted for the entire range of delays between 0 and 100% of the filter length. For the longer filter, this effect is predicted to be larger (1.59 dB). In both cases, the maximal noise reduction is predicted at a delay of 50% of the filter length. The experimental results show an amount of noise reduction and—to a certain degree—a shape of the curves which are similar to the predicted ones. However, there is one major difference: The maximal noise reduction is reached at delays between 12.5% and 25% of the filter length and not at 50%, as predicted. This holds for both the real and the simulated environment and for both filter lengths. For the shorter filter, there is a second, only slightly smaller maximum at a delay of 0% in both environments. In this respect, the theoretical prediction clearly fails. The reasons for and implications of this failure will be discussed in Sec. IV E.

The last design parameter to be considered is the sampling rate F_s in Fig. 13. In real applications, the range of possible values is small, as sampling rates below approximately 7000 Hz will reduce speech recognition unacceptably, and the computational load rises rapidly with higher sampling rates. For the range of $F_s = 5120–20\,480$ Hz, the SNR improvement drops on the order of magnitude of 2 dB. This effect can be explained by the effectively shorter filter (12.5 ms at $F_s = 20\,480$ Hz, compared to 50 ms at $F_s = 5120$ Hz) for the same number of coefficients $N = 256$, which was kept constant.

F. SNR improvement and intelligibility-weighted gain

Figure 14 shows the comparison between SNR improvement (as shown in Figs. 3–13) and intelligibility-weighted gain G_i (Greenberg *et al.*, 1993) for all experimental results which were compared to the theoretical predictions. Although differences up to 2.11 dB do exit, for the majority of all data points SNR improvement and intelligibility-weighted gain G_i are within 0.5 dB. On the average, SNRs are slightly higher than G_i for experiments in real environment (average difference +0.37 dB), while SNRs are slightly lower for the experiments in simulated environments (average difference −0.34 dB).

IV. DISCUSSION

A. Agreement between predictions and measurements

Agreement between experimental results and predictions appear to be reasonable for low direct-to-reverberant ratios of the noise signal $P_{\mathrm{d/r}}$, but considerably poorer for high direct-to-reverberant ratios (cf. Figs. 4 and 9). When the experimental results are compared for all 88 experiments with a $P_{\mathrm{d/r}} < +3$ dB, an average difference of −0.23 dB and a standard deviation of 0.54 dB can be observed. For the 32 experiments in real rooms, the mean difference is −0.21 dB (std. dev. 0.59 dB) and for the 56 experiments in simulated rooms it is −0.25 dB (std. dev. 0.51 dB). As to the 4 experiments with a $P_{\mathrm{d/r}}$ above +3 dB, one comparison with the predicted values is not possible, as the predictions fail at infinite $P_{\mathrm{d/r}}$ (anechoic environment, Fig. 9), and for the other 3 cases differences up to 5.34 dB (Fig. 4) can be observed. From the assumptions of the underlying theoretical analysis (Kompis and Dillier, 2001), it can be expected that agreement between experimental results and predictions is reasonable for low direct-to-reverberant ratios of the noise signal $P_{\mathrm{d/r}}$, but poor for high direct-to-reverberant ratios. From the results shown in Figs. 4 and 9 it can be concluded that the prediction is reasonable for situations with direct-to-reverberant ratios of the noise source of up to approximately +3 dB, while noise suppression is underestimated for higher $P_{\mathrm{d/r}}$. In contrast, the influence of direct-to-reverberant ratio of the *target* signal source appears to be small.

If $P_{\mathrm{d/r}}$ is less than +3 dB, predictions give, on average, a slightly (0.23 dB) higher noise suppression than the experimental results. As the delay in the target-signal path (between d' and d in Fig. 1) is kept at a suboptimal 50% of the filter length (cf. Fig. 12) for the majority of the experiments, it can be expected that experimental results are slightly poorer than the predictions. The standard deviation of the differences between predicted and measured values of approximately 0.5 dB is comparable to the small variations in results, if, e.g., the entire experimental apparatus is shifted by a few centimeters in any direction, as verified by informal tests (Kompis and Dillier, 1993). As seen in Table III, and confirmed by the data presented in Figs. 3–13, results of the experiments in real and simulated rooms are in reasonable agreement, thus supporting the assumption that the chosen room-simulation algorithm (Kompis and Dillier, 1993) is suitable for these experiments involving the adaptive beamformer. One difference between the results of the experiments in real and simulated environments is the tendency to overestimate intelligibility-weighted gain G_i by using SNR improvements for real rooms, and underestimate G_i for simulated environments. This relatively small difference can be attributed to the small differences in the spectra of the simulated and recorded signals.

B. Influence of the noise signal source

From the three parameters defining the noise source, the azimuth has the smallest effect on the amount of noise reduction of the adaptive beamformer. Experimental results and predictions are in reasonable agreement.

Noise reduction is greatly increased with smaller distances between noise source and listener (Fig. 4). The agreement between experimental results and predicted noise reduction is reasonable for distances of 0.75 m or more, but poor for smaller distances. From the assumptions used to derive the predictions (Kompis and Dillier, 2001) it is known that the direct-to-reverberant ratio of the noise signal may not become too large for the predictions to remain valid. The estimated direct-to-reverberant ratio is 2.0 dB at 75 cm and 5.5 dB at 50 cm. According to the presented data, the transition between reasonable prediction and substantial underestimation of the SNR improvement takes place in this range. Although this does limit the usefulness of the prediction method in environments with no or very little reverberation, it is not a serious limitation for most normal rooms with realistic amounts of reverberation. To reach a direct-to-reverberant ratio of +3 dB or more in the room used for the central setting, an omnidirectional noise source must be less than 36 cm away from the listener.

For the investigated range of the index of directionality of the noise source ($\gamma_n = 1–5$), noise reduction changes only moderately (order of magnitude 1–2 dB), depending on the reference signal (left microphone, right microphone, or sum of microphone signals) to which SNR improvement is compared.

C. Influence of the target signal source

The influence of the target signal source on the performance of the adaptive beamformer is small. For the entire range of the alignment factor $A=2$ to $A=4$, the measured and predicted SNR improvement changes by less than 1 dB. SNR improvements differ by the same order of magnitude for the range of values considered for the distance to the signal sound source ($l_s=0.25-1.75\,\mathrm{m}$) as well as the index of directionality ($\gamma_s=0$ to 5).

D. Influence of room size and reverberation

While room size has only a limited effect on the performance of the adaptive beamformer at a fixed reverberation time (Fig. 10), noise reduction drops rapidly with increasing reverberation times T_r (Fig. 9). This phenomenon has been reported previously by several researchers (Peterson *et al.*, 1987; Greenberg and Zurek, 1992; Dillier *et al.*, 1993; van Hoesel and Clark, 1995).

Experimental results and predictions are in reasonable agreement for the range of room sizes considered in this study and for reverberation times of 0.2 s and more, i.e., corresponding to direct-to-reverberant rations $P_{\mathrm{d/r}}$ of the noise source of $+2.7$ dB or less. As noted earlier, predictions systematically underestimate the noise reduction for lower T_r and—consequently—higher $P_{\mathrm{d/r}}$.

E. Influence of the design parameters of the adaptive beamformer

From the three design parameters considered, the sampling rate (Fig. 13) has the smallest range of reasonable values and at the same time a relatively small impact on the performance. The length of the adaptive filter significantly influences the SNR improvement of the adaptive beamformer, as has been noted by several researchers (Peterson *et al.*, 1987; Peterson *et al.*, 1990; Greenberg and Zurek, 1992; Dillier *et al.*, 1993). From our data, we conclude that short filters (e.g., $N=32$) improve SNR to only little above the SNR of the microphone with the more favorable SNR. Only a longer filter in the range of 128–512 coefficients provides substantial additional gains in SNR of 2–4 dB.

The influence of the amount of delay (Fig. 12) is clearly not predicted correctly. For both filter lengths and in both the real and the simulated environment, optimal performance of the beamformer is reached at considerably shorter delays than the predicted 50% of the length of the adaptive filter. This may also explain why in Fig. 11 the agreement between prediction and experimental result is poorest for the longest filter, where the influence of the delay is largest. The reason for the shorter optimal delay is not completely understood. Preliminary results from a small separate investigation suggest a loose relationship between the direct-to-reverberant ratio $P_{\mathrm{d/r}}$ of the noise signal and the optimal delay: for small $P_{\mathrm{d/r}}$, the optimum seems to be close to the predicted 50%, whereas for greater $P_{\mathrm{d/r}}$, e.g., above 0 dB, the optimum tends to be often between 12.5% and 25%.

F. Applicability of the results for hearing aid applications

The presented experiments involve several simplifications, which are not necessarily met in real-life situations encountered by potential future users of an adaptive beamformer. These simplifications, which were made necessary by the assumptions on which the theoretical predictions are based, include: (1) a completely adapted filter, (2) filter adaptation in the absence of the target signal, (3) white noise emitted by both the noise and the target signal source, (4) no movement of either listener or either sound source, and (5) the presence of a single noise source only. Because of the usually fast adaptation time constants [order of magnitude: below 0.1 s (Dillier *et al.*, 1993)] and the availability of several target-signal-detection/adaptation-inhibition schemes (Van Compernolle, 1990; Greenberg and Zurek, 1992; van Hoesel and Clark, 1995; Kompis *et al.*, 1997) assumptions (1) and (2) are likely to be reasonably approached in real-life situations. Acoustic signals in relevant everyday situations will probably be composed of predominately low frequency signals such as speech and traffic noise rather than white noise, as assumed here. However, many implementations of adaptive beamformers use pre-emphasis filters just after the microphones, which prewhiten the spectra of these signals. Therefore, the spectras of probable real-life acoustic signals will at least approach that of white noise to a certain degree. SNR improvements appear to be reasonable estimates for the expected improvement in intelligibility in a number of situations as shown by the data in Fig. 14 and confirmed by tests using a portable real-time realization of the adaptive beamformer (Kompis *et al.*, 1999).

In every-day situations, a certain amount of relative movement between the listener and sound sources must be expected. It is difficult to estimate the influence of such movements. However, due to the usually short adaptation time constants this influence may be small. As to the presence of multiple noise sources, a limited number of experiments have already been reported (Peterson *et al.*, 1990). A substantial drop in performance can be expected if the spectra and levels of the sound sources are similar (Greenberg and Zurek, 1992).

V. SUMMARY AND CONCLUSIONS

A method to predict the amount of noise reduction which can be achieved using a two-microphone adaptive beamforming noise reduction system for hearing aids (Kompis and Dillier, 2001) was verified experimentally. 92 experiments were performed in real and simulated environments and the results were compared with the predictions. It was shown that predictions and experimental results agree reasonably, if the direct-to-reverberant ratio $P_{\mathrm{d/r}}$ of the noise source is smaller than approximately $+3$ dB. For higher $P_{\mathrm{d/r}}$, the predictions systematically underestimate the performance of the adaptive beamformer. The parameters with the greatest influence on the performance of the adaptive beamformer were found to be the direct-to-reverberant ratio of the noise source, the reverberation time of the acoustic environment, and the length of the adaptive filter.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Research Foundation, Grant Nos. 4018-10864 and 3238-56352.99, and Ascom Tech Ltd.

Allen J. B., and Beakley, D. A. (**1979**). ''Image method for efficiently simulating small-room acoustics,'' J. Acoust. Soc. Am. **65**, 943–950.
ANSI (**1969**). ANSI S3.5-1969, ''American National Standard Methods for the Calculation of the Articulation Index'' (American National Standards Institute, New York).
Bellanger, M. G. (**1987**). *Adaptive Digital Filters and Signal Analysis* (Marcel Dekker, New York).
DeBrunner, V. E., and McKinney, E. D. (**1995**). ''Directional adaptive least mean square acoustic array for hearing aid enhancement,'' J. Acoust. Soc. Am. **98**, 437–444.
Dillier, N., Fröhlich, T., Kompis, M., Bögli, H., and Lai, W. L. (**1993**). ''Digital signal processing (DSP) applications for multiband loudness correction digital hearing aids and cochlear implants,'' J. Rehabil. Res. Dev. **24**, 95–109.
Greenberg, J. E., and Zurek, P. M. (**1992**). ''Evaluation of an adaptive beamforming method for hearing aids,'' J. Acoust. Soc. Am. **91**, 1662–1676.
Greenberg, J. E., Peterson, P. M., and Zurek, P. M. (**1993**). ''Intelligibility-weighted measures of speech-to-interference ratio and speech system performance,'' J. Acoust. Soc. Am. **94**, 3009–3010.
Griffiths, L. J., and Jim, C. W. (**1982**). ''An alternative approach to linearly constrained adaptive beamforming,'' IEEE Trans. Antennas Propag. **30**, 27–34.
Hamacher, V., Mauer, G., and Döring, W. H. (**1996**). ''Untersuchung eines adaptiven Beamforming-Systems zur Störunterdrückung für Hörgeschädigte,'' Proceedings of the 22nd Deutsche Jahrestagung für Akustik (DAGA), Bonn, Germany (unpublished).
Kiefer, J. *et al.* (**1996**). ''Speech understanding in quiet and in noise with the CIS speech coding strategy (MED-EL Combi-40) compared to the multipeak and spectral peak strategies (Nucleus),'' ORL **58**, 127–135.
Kochkin, S. (**1993**). ''Consumer satisfaction with hearing instruments in the United States,'' The Marketing Edge, Special issue June 1993, 1–4.
Kompis, M., and Dillier, N. (**1993**). ''Simulating transfer functions in a reverberant room including source directivity and head shadow effects,'' J. Acoust. Soc. Am. **93**, 2779–2787.
Kompis, M., and Dillier, N. (**2001**). ''Performance of a two-microphone adaptive beamforming noise reduction scheme for hearing aids. II. Experimental verification of the predictions,'' J. Acoust. Soc. Am. **109**, 1123–1133.
Kompis, M., Dillier, N., Francois, J., Tinembart, J., and Häusler, R. (**1997**). ''New target-signal-detection schemes for multi-microphone noise-reduction systems for hearing aids,'' Proc. Annu. Int. Conf. IEEE Eng. Biol. Soc. **19**, 1990–1993.
Kompis, M., Feuz, P., François, J., and Tinembart, J. (**1999**). ''Multi-microphone digital-signal-processing system for research into noise reduction for hearing aids,'' Innovation Technol. Biol. Med. **20**, 201–206.
Kuhn, G. F. (**1977**). ''Model for the interaural time difference for the azimuthal plane,'' J. Acoust. Soc. Am. **62**, 157–167.
Peterson, P. M., Durlach, N. I., Rabinowitz, W. M., and Zurek, P. M. (**1987**). ''Multimicrophone adaptive beamforming for interference reduction in hearing aids,'' J. Rehabil. Res. Dev. **24**, 103–110.
Peterson, P. M., Wie, S. M., Rabinowitz, W. M., and Zurek, P. M. (**1990**). ''Robustness of an adaptive beamforming method for hearing aids,'' Acta Oto-Laryngol., Suppl. **469**, 85–90.
Van Compernolle, D. (**1990**). ''Hearing aids using binaural processing principles,'' Acta Oto-Laryngol., Suppl. **469**, 76–84.
van Hoesel, R. J. M., and Clark, G. M. (**1995**). ''Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients,'' J. Acoust. Soc. Am. **97**, 2498–2503.
Widrow, B., Glover, J. R., McColl, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, J. R., and Goodlin, R. C. (**1975**). ''Adaptive noise cancelling: Principles and applications,'' Proc. IEEE **63**, 1692–1716.