Development and assessment of two fixed-array microphones for use with hearing aids

Prof. Dr. Frans A. Bilsen; Dr. Wim Soede; Prof. Dr. Augustinus J. Berkhout
Delft University of Technology, Faculty of Applied Physics, 2600 GA Delft, The Netherlands

Abstract—Hearing-impaired listeners often have great difficulty understanding speech in situations with background noise (e.g., meetings, parties). Conventional hearing aids offer insufficient directivity to significantly reduce background noise relative to the desired speech signal. Based on array techniques, microphone prototypes have been developed with strongly directional characteristics to be incorporated into the frame and the “temples” of a pair of eyeglasses. Particular emphasis was on optimization and electronic stability. Computer simulations show that a directivity index of more than 10 dB can be obtained at the higher frequencies. Simulations were verified with free-field measurements. To investigate the influence of the human head on directivity, two portable models were also tested with a KEMAR manikin. The measurements show that the two models give an improvement of the signal-to-noise ratio of approximately 7 dB in a diffuse background noise field compared with an omnidirectional microphone. For the clinical assessment of these microphone arrays in the diffuse noise field (simulating a cocktail party situation), the speech-reception threshold in noise for simple Dutch sentences was determined with a normal single omnidirectional microphone and with one of the microphone arrays. The results of monaural listening tests of 30 subjects with normal hearing and 45 subjects with hearing impairment show that the microphone arrays give a mean improvement of the speech reception threshold in noise of about 7 dB compared with an omnidirectional microphone.

Key words: background noise, KEMAR manikin, microphone arrays, microphone eyeglass prototypes, omnidirectional microphone, signal-to-noise ratios, speech reception.

INTRODUCTION

Many people have great difficulty understanding speech in surroundings with background noise and/or reverberation. This is especially a problem for the increasing number of elderly people and people with sensorineural impairment. Several investigations of speech intelligibility in noisy situations have demonstrated that subjects with sensorineural hearing loss may need a 5–15 dB higher signal-to-noise ratio than subjects with normal hearing (1). Every 4–5 dB improvement of the signal-to-noise ratio may raise the speech intelligibility by about 50 percent (2,3,4).

A directional hearing aid may reduce background noise relative to the desired speech signal. Until now, directional hearing aids consisted of a conventional hearing aid with a single cardioid microphone. Although Mueller (5) and Hillman (6) published studies showing a preference for a hearing aid with a directional cardioid microphone, the directional hearing aid has not yet enjoyed the widespread clinical acceptance that would be expected on theoretical grounds. In practice, the reduction of background noise with a cardioid microphone is not yet sufficient because the low directivity for higher frequencies (2,000–5,000 Hz) permits a maximum improvement of only about 2 dB (7,8).
Different solutions to further improve the directivity seem plausible on theoretical grounds. In adaptive processing, the processing of the signals from two or more microphones is continuously adjusted according to properties of the received sound signals and controlled by an adaptation mechanism that can be implemented on a signal processor (8,9,10,11).

In fixed array processing, a fixed configuration of a number of microphones offers the possibility of suppression of background noise while the desired speech signal in front of the user is transmitted undistorted. A high and robust directivity can be obtained when the length of the array is larger than the wavelength; the signal processing can be done with relatively simple analogue electronics. A desired sound source can be chosen by moving the principle direction of the array toward the source. A cosmetic disadvantage of a fixed array technique might be the array length. As a practical compromise, it was envisaged that the microphone array should be connected to, or built into, a pair of eyeglasses and should be used in combination with a conventional hearing aid. Therefore, the maximum array length is determined by the length of the eyeglass "temples" (i.e., the pieces that extend from the frames alongside the head and around the ears) or the width of the frame, which is approximately 10 cm and 14 cm, respectively.

The directivity index (DI) was accepted as a measure to differentiate between possible solutions. It was decided to optimize the DI within a frequency range of 500 Hz to 4,000 Hz. The shape of the directivity pattern was considered to be of secondary importance. Further, the new directional microphone is meant to be used monaurally. Profits of binaural fitting should be added by simply using two devices.

In the following sections, the results of computer simulations and measurements on different array configurations, optimization and stability of different array configurations, and listening tests with normal-hearing and hearing-impaired subjects will be summarized (12).

METHODS AND DISCUSSION

Simulations on Broadside and Endfire Array

For an application with microphones mounted on a pair of eyeglasses, we distinguish between two important groups of linear arrays characterized by the position of the microphones, viz. broadside arrays and endfire arrays.

In a broadside array the microphones are placed along the x-axis (alongside each other). The directivity pattern can be shown (13) to be given by

$$Q(\theta, \phi, \omega) = \sum_n D_n(\omega)A_n(\omega)e^{+jk(\omega)n\Delta x}$$ \[1\]

with $\theta$ and $\phi$ the angles of incidence, $\omega$ the frequency, $n$ the microphone number, $D_n(\omega)$ the frequency-dependent directivity characteristic of the individual microphones, $A_n(\omega)$ the amplitude weighting of individual microphones, $k_x$ equal to $k \cos \phi \sin \theta$ ($k$ is the wave number), and $\Delta x$ the distance between the microphones.

In an endfire array the microphones are placed along the z-axis (behind each other) and the phase correction should be $\tau_n = n\Delta z/c$ ($c$ is sound velocity). Now, the directivity pattern is given by

$$Q(\theta, \phi, \omega) = \sum_n D_n(\omega)A_n(\omega)e^{-j\omega\tau_n}e^{jk(\omega)n\Delta z}$$ \[2\]

with $k_z = k \cos \theta$.

Considering the situation with the desired sound coming from the main direction of the array and the background noise distributed equally over all other directions, the directivity index $DI(\omega)$ is a proper measure to indicate the average attenuation of the background noise with respect to sound coming from the main direction (14); it is given by

$$DI(\omega) = 10 \log \left( \frac{4\pi |Q(\theta, \phi, \omega)_{max}|^2}{\int_{0}^{2\pi} \int_{0}^{\pi} |Q(\theta, \phi, \omega)|^2 \sin \theta \, d\theta \, d\phi} \right)$$ \[3\]

For a broadside array of five microphones with total length $L = n\Delta x = 10$ cm assuming omnidirectional microphones [$D_n(\theta, \phi, \omega) = 1$] and uniform amplitude weighting ($A_n = 1$), the DI equals 4.9 dB at 4,000 Hz. An endfire array with similar parameters will have a DI = 7.6 dB.

A comparison of equations for the beam width of the broadside and endfire array shows that in the x-z plane the main beam of a broadside array is always narrower than that of an endfire array of the same size. Therefore, the use of a broadside array may be advantageous when a small beam width is wanted in one plane. However, the computed directivity indices show that an endfire array of
omnidirectional microphones is advantageous for diffuse noise suppression.

Amplitude weighting \( [A_n(\omega)] \) of each microphone signal is equivalent with the application of a window function. The uniform weighting gives a small main lobe with relative high side lobe levels. A concave upward weighting results in a narrower beam width at the expense of having higher side lobe levels. The opposite effect can be obtained with a Cosine, Hanning or Dolph-Chebyshev window function. They reduce the side lobe levels at the cost of a broader main lobe and a lower DI. The broader main lobe is a result of the low amplitude weighting at both ends of the array, giving an array with a relatively shorter effective length. However, the uniform weighting and the concave upward weighting have the highest DI. Finally, it must be noted that the weighting functions can also be applied to an endfire array. The amplitude weighting is independent of the phase correction, but both can be used to shape the directivity pattern (12).

Using cardioid microphones \( [D_n(\theta,\phi,\omega) \neq 1] \) can be very useful in array design to improve the directivity for the lower frequencies (\( \lambda > L \)), to suppress side lobes and/or unwanted main lobes. This is especially advantageous for low frequencies and for suppressing the backward lobe of the broadside array.

Figure 1 gives the directivity pattern and the directivity index at 4,000 Hz for both array configurations with five cardioid microphones in a free-field situation, for a broadside array with \( L = 14 \) cm and an endfire array with \( L = 10 \) cm. The cardioid microphones give a significant improvement of the DI at the lower frequencies. We may conclude that with one endfire array (\( L = 10 \) cm) or one broadside array (\( L = 14 \) cm) of cardioid microphones, a DI can be reached of at least 5 dB at the lower frequencies rising to more than 10 dB at 4 kHz.

Combinations can be made of one endfire and one broadside array, two endfire arrays with intermediate distance \( L = 14 \) cm, and a full configuration of two endfire arrays and one broadside array having the shape of one pair of eyeglasses. The directivity of the combined configurations was computed for a free-field situation and a simple summation of the (delayed) microphone signals giving one output signal (mono). The combined array configurations (mono, free-field) give an extra improvement of the DI between 2 and 3 dB.

**A Numerical Approach to Optimization and Stability**

For an array configuration to be optimally effective in the improvement of speech intelligibility in noise, it should have optimal directional characteristics. But an array configuration with a high directivity may not necessarily be a stable one. The high value of the DI may be reached for a special theoretical parameter choice that can be created in a laboratory situation but cannot be maintained in practice. Therefore, in a practical situation it is important to choose a stable array solution with a sufficient directivity, which is minimally influenced by intrinsic variabilities of microphones (amplitude and phase characteristics), amplitude weighting (amplifiers and resistor values), and/or time lag correction (delay elements).

Soede (12) used a comprehensive quasi-Newtonian algorithm (15) to study endfire, broadside, and Jacobi arrays. The algorithm searches for an unconstrained maximum of a function vector \( F \)
of parameters (represented by parameter vector x), where no mathematical derivatives of the function are required. The variables can be subjected to fixed lower and/or upper bounds. In this application the function vector F was defined by the equation for the DI. The optimization process was executed for single frequencies (i.e., 1,000, 2,000, 3,000, and 4,000 Hz) with respect to the amplitude weighting and the time lag correction of each cardioid microphone. The optimization processes showed that optimization of the parameter set at 4,000 Hz was sufficient for this application, giving a high DI for the lower frequencies too.

For the endfire array, optimization was done with a fixed overall length of 10 cm and for a changing number of microphones ranging from 2 to 17 with optimal time lags according to Hansen and Woodyard (16). It turned out that the DI at 4 kHz improves by 4 dB when five or more microphones are used instead of one. Using six or more microphones gives no further improvement. With respect to stability, it was shown that the influence of variations in amplitude weighting and delay times on the value of the DI of an optimized endfire array with five microphones is small. Therefore, it was concluded that an optimized endfire array with five microphones offers a stable and practical solution.

For the broadside array, optimization can be performed with respect to the position of the microphones as well as their amplitude weighting. Regarding the first, Ma (17) showed that the directivity of an array with variable microphone spacing will, at its optimum, be only a fraction of a decibel higher than an array with equidistant microphone spacing. Therefore, Soede (12) only paid attention to amplitude weighting with equidistant microphones. The optimization was done for a broadside array with a fixed width of 14 cm consisting of four, five, and six microphones. It was found that the profit of the optimization is very small. The directivity is mainly determined by the length of the array in relation to the wave length. Also, the spectrum of the optimized broadside array is hardly influenced by the optimization and is equal to the spectrum of one single cardioid microphone. A comparison of DIs for optimized configurations is presented in Figure 2.

With respect to stability, the influence of variations in the amplitude weighting appears to be very small as well. A difference of 0.2 dB between the DI of the uniform broadside array and the optimized array was reached by a variation in the amplitude weighting of more than 50 percent. Thus, variations with respect to sensitivity between the microphones used are negligible. Variations with respect to phase characteristics, on the other hand, reduce the directivity for a frequency of 4,000 Hz, but not below 6 dB. In summary, up to 2,000 Hz a broadside array appears very stable with respect to variations in individual microphone components—for further details see (12).

**Directivity Measurements in Free Field and with KEMAR**

For an assessment of the microphone arrays, a laboratory model of an endfire array and a broadside array was built (12). Because it was expected that the directivity of an array might be influenced by reflections and diffractions at the head, measurements were carried out in an anechoic chamber with the models placed in free-field conditions as well as in combination with an artificial head (KEMAR).

The laboratory model consisted of directional electret microphones (MICROTEL 61) with tube extensions to obtain a cardioid directivity pattern. The microphones were connected to movable little sockets. In the endfire configuration, each microphone is placed with its maximum sensitivity to \( \theta = 0^\circ \) (along the bar). The signal of each microphone is delayed relative to the first microphone signal using Panasonic MN3012 bucket-delays. The delay time of each microphone could be varied.
independently of the other delays by variation of the clock frequency of the bucket-delay. Amplitude weighting is done with adjustable amplifiers. For the broadside array, the same laboratory model with delay times set to zero was used. Each microphone was placed in the broadside configuration with its maximum sensitivity to $\theta = 0^\circ$ (perpendicular to the bar).

Directivity patterns were measured monochromatically in an anechoic chamber ($V = 1,000 \text{ m}^3$) with the microphone array mounted on a turntable and with a loudspeaker at a distance of 6.4 m. Data acquisition was carried out with a PC-controlled measurement system developed in-house. For practical reasons, the directivity patterns were measured in the horizontal plane only. For the endfire array, the DI was computed from the measured directivity pattern assuming the main beam at $0^\circ$ and a cylinder symmetry along that main beam. For the broadside array, the DI was computed from the horizontal directivity pattern and corrected for the cardioid-like directivity pattern in the vertical plane.

For an optimized endfire array with five cardioid microphones and a length of 10 cm, directivity patterns (polar diagrams) are presented in Figure 3 for free-field conditions (a) and with KEMAR (b). Directivity indices were computed from these measurements with the restrictions mentioned in the previous paragraph. A comparison of the free-field directivity patterns and those measured with KEMAR show that, especially when the sound is coming from the right side of KEMAR, the influence of the head on the performance of the array is relatively small. The reduction of the estimated DI is less than 1 dB for all frequencies. The main beam is in the direction of $0^\circ$ for all frequencies. Apparently, the directivity of the endfire array is hardly decreased by the addition of reflections or diffraction of the sound by the head.

For a broadside array with five cardioid microphones and total width of 14 cm, directivity patterns are given in Figure 4 for free field conditions (a) and with KEMAR (b). A comparison shows that the influence of the head on the directivity patterns is again very small and is even advantageous for suppression of the backside lobes. The decrease in the DI at 500 Hz is less than 0.5 dB.

In summary, the estimated values of the DI computed from the polar patterns measured with the KEMAR manikin show that one cardioid microphone may give a mean improvement of 4 dB in comparison with one omnidirectional microphone. The estimated DI of the optimal endfire microphone array as well as the broadside microphone array ranges from 4 dB at 500 Hz to more than 10 dB at 4,000 Hz.

**Directivity Measurements in an Artificial Diffuse Noise Field**

Based on the results summarized in the previous sections, two portable array models were built that were suitable for psychophysical assessment with hearing impaired listeners: a portable endfire microphone array with a total length of 10 cm, and a portable broadside microphone array with a total width of 14 cm mounted on a pair of eyeglasses (12). Because cardioid microphones have a spectrum rising with 6 dB/octave, a correction filter for flattening was applied. The correction filter was designed in such a way that the spectrum in the front direction of both arrays was flat within ±2 dB between 500 and 4,000 Hz.
An artificial diffuse sound field mimicking a cocktail party situation was realized with eight small loudspeakers positioned at the boundaries of an imaginary rectangular box (2.0 × 2.0 × 1.70 m) inside a sound-insulated booth (at the ENT department of the University Hospital, Rotterdam). Four loudspeakers were placed near the ceiling of the room at the edges of the rectangular box and the other ones were placed vertically (height 40 cm) at the corners of the cube. The eight loudspeakers were fed with eight independent noise sources, producing a spectrum equal to the long-time-average spectrum of speech. One loudspeaker was positioned at a height of 1.25 m in front of the listener (seated near the center of the box) and simulated the partner in a discussion during the listening test. The sound levels of the speech and the noise field could be varied with an audiometer and an 8-channel attenuator developed in-house, respectively; both variables were controlled by a personal computer—for further details, see (12).

A KEMAR manikin was placed in the center of the experimental set-up facing the front loudspeaker (distance 1 m), and measurements were carried out with two behind-the-ear hearing aids, one with an omnidirectional microphone and one with a cardioid microphone, and then with the portable broadside and endfire microphone arrays. The hearing aids were connected to the right ear of KEMAR with an earmold (libby-horn with foam plug). Signals of the microphone arrays were measured via the behind-the-ear hearing aid using an induction-loop and the induction coil of the hearing aid (this equals the listening test conditions with the real subjects). With this set-up the attenuation of the diffuse noise field relative to the noise coming from the front direction was measured in one-third-octave bands.

The results of the measurements are reproduced in Figure 5. The attenuation is given as a function of the one-third-octave center frequency (400–5,000 Hz). The mean level is computed from the one-third-octave band levels with equal weights. The measurement with the hearing aid containing a normal omnidirectional microphone shows that the diffuse sound field is not attenuated. The mean value of –1 dB means an amplification of the diffuse sound field relative to the signal coming from the front. The hearing aid with one cardioid microphone attenuates the diffuse sound field for the lower frequencies with a mean of +1.5 dB, indicating an improvement of +2.5 dB compared with the omnidirectional hearing aid. This corresponds with everyday experience (5,6).

The measurements with the broadside and endfire microphone arrays show a strong attenuation, especially for the high frequencies, with mean values of +6.0 dB and +5.8 dB, respectively, thus indicating an improvement of 7.0 dB (broadside) and 6.8 dB (endfire) compared with the omnidirectional hearing aid. These results are about 1 dB lower in comparison with the DIs estimated from the KEMAR directivity patterns. This difference can be explained by a contribution of the sound of the front loudspeaker to the diffuse noise field due to (not negligible) reverberation in the soundproof booth.

**Figure 4.**
Measured polar diagrams of an optimized broadside array consisting of five cardioid microphones and an overall width of 14 cm, in free-field (a) and with KEMAR (b).
Psychophysical Assessment in a Cocktail-Party Simulation

Speech intelligibility with the arrays in a cocktail-party situation was determined using the artificial diffuse noise field as briefly described in the previous section. Using a simple up-down procedure, the 50 percent intelligibility level, the so-called speech reception threshold (SRT) was determined. The difference between the SRT in noise and the noise level was defined as the critical speech-to-noise ratio (S/N ratio). The speech material consisted of 10 lists of 13 short Dutch sentences, representative of everyday conversation (4). For the simulation of the background noise at a typical cocktail party, eight independent "speech noise" signals were used having a spectrum equal to the long-term-average spectrum of the sentences.

The monaural listening tests were carried out with 30 subjects with normal hearing and 45 subjects with hearing impairment. The group of 30 normal-hearing listeners, equally divided as to males and females, were mainly physics and medicine students with ages ranging from 19 to 37 years with a median age of 26.4 years. The hearing-impaired group consisted of 23 male and 22 female subjects, aged 36–90 years, with a median age of 68.8. The hearing-impaired listeners were asked for their cooperation while visiting the ENT department of the Rotterdam University Hospital for hearing-aid inspection. Cooperation was requested after the hearing-aid fitting received its final approval and when a discrimination score of at least 80 percent for monosyllables presented in quiet was found. Each listening test took about 15 minutes.

For the assessment of the microphone arrays with hearing-impaired subjects, an induction loop was used in combination with the subject's own individually fitted hearing aid—for experimental details see (12). If necessary, the other ear was occluded with a foam plug (E.A.R.-plug). Because we were primarily interested in comparing the omnidirectional microphone and the microphone arrays, most subjects with hearing impairment performed two listening tests under two conditions: with their own hearing aid in combination with an external omnidirectional microphone, and their own cardioid, the broadside, and the endfire microphone. The mean level is computed from the 1/3-octave measurements with equal weights.

Figure 5.
Attenuation of the artificial diffuse sound field, as measured in 1/3-octave bands with KEMAR and an omnidirectional, a

...
hearing aid with endfire or broadside microphone array. A subgroup of eight hearing-impaired listeners took the listening tests under all three conditions. The order of conditions (e.g., with and without microphone array) was varied to avoid effects of habituation and fatigue, and, moreover, to equalize small differences between the 10 lists of sentences.

For the group of normal-hearing subjects, the listening tests were carried out with changing subgroups for four conditions: own ear, hearing aid (Philips M47), broadside microphone array, and endfire microphone array via hearing aid. The normal-hearing subjects listened to the hearing-aid using a libby-horn with foam plug.

Figure 6 presents the averaged S/N ratios and intersubject standard deviation for the number of listening tests (n) per condition, for the normal-hearing group as well as for the hearing-impaired group. In addition, the values of the S/N ratios at overall levels of +20 and +30 dB and the difference between these values are given. The small differences between the values of the S/N ratios at +20 and +30 dB confirm the linearity of the SRTs as hypothesized by Plomp (3) and the reliability of the mean S/N ratios for most conditions. For the normal-hearing conditions, the standard deviations are less than 1.2 dB. For the hearing-impaired conditions, they are about 3 dB.

A comparison of the S/N ratios shows the following points:

1. The monaural S/N ratio of the normal-hearing group (listening with one good ear) equals -8.5 dB.
2. A hearing aid (Philips M47) decreased the S/N ratio of the normal-hearing listeners by 1.2 dB.
3. The S/N ratio of the normal-hearing group can be improved significantly using a microphone array, instead of an omnidirectional microphone.
4. The hearing-impaired group listening with the omnidirectional microphone has a S/N ratio of only -0.2 dB, with a large standard deviation of 3.4 dB.
5. The microphone arrays also give a significant improvement of the S/N ratios for the hearing-impaired group. The absolute values of the average S/N ratio obtained with the microphone arrays is comparable with the S/N ratio of the normal-hearing group listening with one good ear.

On the average, the broadside microphone array gives an improvement of 7.0 dB with a standard deviation of 1.9 dB, and the endfire microphone array gives an improvement of 6.8 dB with a standard deviation of 2.1 dB. Most hearing-impaired listeners (41 of 45) obtain an improvement of at least 5 dB. The difference of 0.2 dB between the broadside and endfire microphone arrays was also found for the subgroup of eight subjects listening to both microphone arrays.

**CONCLUSION**

The following was noted as the result of the work described above:

- Computer simulations showed that it should be possible to construct a broadside or an endfire array with dimensions suitable for mounting on a pair of eyeglasses, with a stable and near-to-optimal directivity index using five cardioid microphones only.
- Directivity pattern and directivity index measured in free-field with and without KEMAR show only a slight degradation in the performance of the arrays due to reflection and diffraction of sound at the head of the order of 1 dB.
- KEMAR measurements in an artificial diffuse noise field show that the broadside microphone array and the endfire microphone array will attenuate the diffuse noise field ("speech noise") relative to sound coming from the front direction ("desired speech") and give a mean improvement of 7.0 dB and 6.8 dB, respectively.
• Speech intelligibility tests in a cocktail party situation (simulated by the diffuse noise field) with normal-hearing and hearing-impaired subjects showed that the microphone arrays improve the critical S/N ratio significantly: an omnidirectional hearing aid microphone broadside and endfire microphone array give a mean improvement of 7.0 dB and 6.8 dB, respectively. These results confirm the KEMAR measurements.

ACKNOWLEDGMENTS

Thanks are due to Dr. J. Verschuure and the members of the ENT department of the University Hospital Rotterdam for their cooperation. This investigation was supported by the Dutch Foundation for Fundamental Research on Matter (FOM) and the Dutch Foundation for Technology (STW).

REFERENCES