Multi-microphone adaptive noise reduction strategies for coordinated stimulation in bilateral cochlear implant devices

Kostas Kokkinakis and Philipos C. Loizou[a]

Department of Electrical Engineering, University of Texas at Dallas, Richardson, Texas 75080

Bilateral cochlear implant (BI-CI) recipients achieve high word recognition scores in quiet listening conditions. Still, there is a substantial drop in speech recognition performance when there is reverberation and more than one interferers. BI-CI users utilize information from just two directional microphones placed on opposite sides of the head in a so-called independent stimulation mode. To enhance the ability of BI-CI users to communicate in noise, the use of two computationally inexpensive multi-microphone adaptive noise reduction strategies exploiting information simultaneously collected by the microphones associated with two behind-the-ear (BTE) processors (one per ear) is proposed. To this end, as many as four microphones are employed (two omni-directional and two directional) in each of the two BTE processors (one per ear). In the proposed two-microphone binaural strategies, all four microphones (two behind each ear) are being used in a coordinated stimulation mode. The hypothesis is that such strategies combine spatial information from all microphones to form a better representation of the target than that made available with only a single input. Speech intelligibility is assessed in BI-CI listeners using IEEE sentences corrupted by up to three steady speech-shaped noise sources. Results indicate that multi-microphone strategies improve speech understanding in single- and multi-noise source scenarios. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3372727]

I. INTRODUCTION

Most cochlear implant (CI) recipients perform well in quiet listening conditions and many users can now achieve word recognition scores of 80% or higher regardless of the device used (Spahr and Dorman, 2004). However, several studies have demonstrated that the ability of CI recipients to correctly identify speech degrades sharply in the presence of background noise and other interfering sounds, when compared against that of normal-hearing listeners (Qin and Oxenham, 2003; Stickney *et al.*, 2004). In an effort to improve speech understanding in noise, individuals with severe to profound hearing loss are now receiving two cochlear implants, one in each ear. Recent clinical studies focusing on assessing overall speech perception in noise with bilateral cochlear implants (BI-CIs) have demonstrated a substantial increase in word recognition performance when compared to monaural listening configurations (e.g., see van Hoesel and Clark, 1997; Lawson *et al.*, 1998; Müller *et al.*, 2002; Tyler *et al.*, 2002; van Hoesel and Tyler, 2003; van Hoesel, 2004; Tyler *et al.*, 2003; Ricketts *et al.*, 2006; Loizou *et al.*, 2009). Large gains in speech reception were reported for both bilaterally implanted adults and children, in quiet and noisy settings (van Hoesel and Tyler, 2003; Litovsky *et al.*, 2004; Grantham *et al.*, 2007; Litovsky *et al.*, 2009). However, under more challenging listening conditions, such as when the noise level is high and the number of interfering maskers is large, BI-CI recipients perform significantly worse than normal-hearing listeners. Recently, speech intelligibility in a multi-source environment, where multiple noise sources were emanating from various directions, was evaluated in terms of the decibel reduction in speech reception threshold.[1] The speech reception threshold (SRT) values obtained by BI-CI users were found to be significantly worse than those obtained by normal-hearing listeners in the same listening conditions (Loizou *et al.*, 2009).

A number of single-microphone noise reduction techniques have been proposed over the years to improve speech recognition in noisy background conditions (Hochberg *et al.*, 1992; Weiss, 1993; Müller-Deile *et al.*, 1995; Yang and Fu, 2005; Loizou *et al.*, 2005; Hu *et al.*, 2007). Modest, but significant, improvements in intelligibility were reported with single-microphone techniques (Yang and Fu, 2005; Loizou *et al.*, 2005; Hu *et al.*, 2007). Considerably larger benefits in speech intelligibility can be obtained when resorting to multi-microphone adaptive signal processing strategies, instead. Such strategies make use of spatial information due to the relative position of the emanating sounds and can therefore better exploit situations in which the target and masker are spatially separated (Kompis and Dillier, 1994; van Hoesel and Clark, 1995; Hamacher *et al.*, 1997; Wouters and Vanden Berghe, 2001; Chung *et al.*, 2006; Spriet *et al.*, 2007).

Nowadays, most bilateral cochlear implant devices are fitted with either two microphones in each ear (e.g., Freedom processor) or one microphone in each of the two (one per ear) behind-the-ear (BTE) processors. The Freedom processor, for instance, employs a rear omni-directional microphone, which is equally sensitive to sounds from all directions, as well as an extra directional microphone pointing forward. The use of directional microphones provides an ef-

---

[a]Author to whom correspondence should be addressed. Electronic mail: loizou@utdallas.edu

fective yet simple form of spatial processing. A number of recent studies have shown that the overall improvement provided by the use of an additional directional microphone can be approximately 3–5 dB in real-world environments with relatively low reverberant characteristics when compared to processing with just an omni-directional microphone (Soede et al., 1993; Wouters and Vanden Berghe, 2001; Chung et al., 2006).

Adaptive beamformers can be considered an extension of differential microphone arrays, where the suppression of interferers is carried out by adaptive filtering of the microphone signals. An attractive realization of adaptive beamformers is the generalized sidelobe canceller structure (Griffiths and Jim, 1982). To evaluate the benefit of noise reduction for CI users, van Hoesel and Clark (1995) tested a two-microphone noise reduction technique, based on adaptive beamforming by mounting a single directional microphone behind each ear. Results indicated large improvements in speech intelligibility for all CI subjects, compared to an alternative two-microphone strategy, in which the inputs to the two-microphone signals were simply added together.

The performance of beamforming algorithms in various everyday-life noise conditions using BI-CI users was also assessed by Hamacher et al. (1997). The mean benefit obtained in terms of the speech reception threshold with the beamforming algorithms for four BI-CI users varied between 6.1 dB for meeting room conditions and just 1.1 dB for cafeteria noise. In another study, Chung et al. (2006) conducted experiments to investigate whether directional microphones and adaptive multi-channel noise reduction algorithms could enhance overall CI performance. The results indicated that directional microphones can provide an average improvement of around 3.5 dB. An additional improvement of 2.0 dB was observed when processing the noisy stimuli through a directional microphone first and then through the noise reduction algorithm, although results were not statistically significant for all conditions tested.

Wouters and Vanden Berghe (2001) assessed speech recognition of four adult CI users utilizing a two-microphone adaptive filtering beamformer. Monosyllabic words and numbers were presented at 0° azimuth at 55, 60, and 65 dB sound pressure level (SPL) in quiet and noise with the beamformer inactive and active. Speech-weighted noise was presented at a constant level of 60 dB SPL from a source located at 90° azimuth on the right side of the user. Word recognition in noise was significantly better for all presentation levels with the beamformer active, showing an average signal-to-noise ratio (SNR) improvement of more than 10 dB. Number recognition in noise was also significantly better with the beamformer active, demonstrating an average SNR improvement of 7.2 dB across conditions. More recently, Spriet et al. (2007) investigated the performance of the BEAM preprocessing strategy in the Nucleus Freedom speech processor with five CI users. The performance with the BEAM strategy was evaluated at two noise levels and with two types of noise, speech-weighted noise and multi-talker babble. On average, the algorithm tested improved (lowered) the SRT by

approximately 5–8 dB, as opposed to just using a single directional microphone to increase the direction-dependent gain of the target source.

It is clear from the above studies that processing strategies based on beamforming can yield substantial benefits in speech intelligibility for cochlear implant users, especially in situations where the target and masker sound sources are spatially separated. Nevertheless, the effectiveness of beamforming noise reduction strategies is limited to (1) only zero-to-moderate reverberation settings (e.g., see Greenberg and Zurek, 1992; Hamacher et al., 1997; Kompis and Dillier, 2001) and (2) single interfering sources. The reverberation time and the direct-to-reverberant ratio[2] have been found to have the strongest impact on the performance of adaptive beamformers (Kompis and Dillier, 1994). The presence of multiple noise sources considerably reduces the overall efficiency and performance of beamforming algorithms. A substantial drop in SRT performance in the presence of uncorrelated noise sources originating from various azimuths has been noted by both Spriet et al. (2007) and Van den Bogaert et al. (2009).

The majority of the above cited studies evaluated beamforming algorithms in situations where a single interfering source was present, and the room acoustics were characterized with low or no reverberation (anechoic settings). In realistic scenarios, however, rooms might have moderate to high reverberation and multiple noise sources might be present, and in some instances, these sources might be emanating from both hemifields (left and right) of the user. In the present study, we take the first step in the development of multi-microphone algorithms, which can be used in realistic scenarios. To that end, we focus on algorithms that can better utilize the information captured by the four microphones (two in each ear) presently available in the Nucleus Freedom processor. Currently, in the Nucleus Freedom processor, two BEAMs (one in each ear) run independently of one another and as such do not utilize intelligently the four microphones available. In situations, for instance, where multiple sources are located in one side, the BEAM algorithm running on the contralateral side could make use of the noise sources impinging on the opposite ear to form a better reference signal.

In the proposed system, we will investigate the potential of contralateral routing of the four-microphone signals with the intention of obtaining a better noise reference signal for the adaptive algorithms. More precisely, we will investigate a new computationally inexpensive multi-microphone noise reduction strategy that can exploit not just two but four microphones. Such a processing strategy can make use of all microphones available to the bilateral users wearing a Nucleus Freedom device in a coordinated fashion and as such, it can better exploit noise sources which are located contralaterally. The idea of exploiting contralaterally placed microphones has been investigated extensively in the hearing aid industry with the use of contralateral routing of signals (Harford and Barry, 1965; Ericson et al., 1988). However, this idea has not yet been explored in the field of cochlear implants. Furthermore, noise reduction strategies that make

use of more than two microphones have never been proposed, implemented, or systematically evaluated with the Nucleus Freedom device.

The underlying hypothesis in this study is that by exploiting the contralateral microphone signals, we can facilitate, or perhaps enhance, the better-ear advantage, which is the most robust bilateral benefit afforded by bilateral implants (Loizou *et al.*, 2009). To test this hypothesis, we will evaluate and compare the performance of two independent BEAMs, currently implemented in the Nucleus Freedom processor, against the proposed strategy that makes use of contralateral signals. Stimuli processed in single- and multi-source reverberant scenarios will be used in the evaluation.

II. SIGNAL PROCESSING STRATEGIES

A. 2M2-BEAM strategy

The 2M2-BEAM noise reduction strategy utilizes the dual-microphone BEAM running in an independent stimulation mode in each of the two (one per ear) BTE units furnished with one directional and one omni-directional microphone, and is therefore capable of delivering the signals bilaterally to the CI user. Each BEAM combines a directional microphone with an extra omni-directional microphone placed closed together in an endfire array configuration to form the target and noise references. The inter-microphone distance is fixed at 18 mm (Patrick *et al.*, 2006). In the BEAM strategy, the first stage utilizes spatial preprocessing through a single-channel, adaptive dual-microphone system that combines the front directional microphone and a rear omni-directional microphone to separate speech from noise. The output from the rear omni-directional microphone is filtered through a fixed finite impulse response (FIR) filter. The output of the FIR filter is then subtracted from an electronically delayed version of the output from the front directional microphone to create the noise reference (Wouters and Vanden Berghe, 2001; Spriet *et al.*, 2007). The filtered signal from the omni-directional microphone is then added to the delayed signal from the directional microphone to create the speech reference. This spatial preprocessing increases sensitivity to sounds arriving from the front while suppressing sounds that arrive from the sides. The two signals containing the speech and noise reference are then fed to an adaptive filter, which is updated with the normalized least-mean-squares algorithm (Haykin, 1996) in such a way as to minimize the power of the output error (Greenberg and Zurek, 1992). The 2M2-BEAM processing strategy is currently implemented in commercial bilateral cochlear implant processors, such as the Nucleus Freedom processor.

B. 4M-SESS strategy

The four-microphone spatial enhancement via source separation (4M-SESS) processing strategy is the two-stage multi-microphone extension of the conventional BEAM noise reduction approach. It is based on coordinated or cross-side stimulation by collectively using information available on both the left and right sides. The 4M-SESS can be classified as a four-microphone binaural strategy, since it relies on having access to four microphones, namely, two omni-
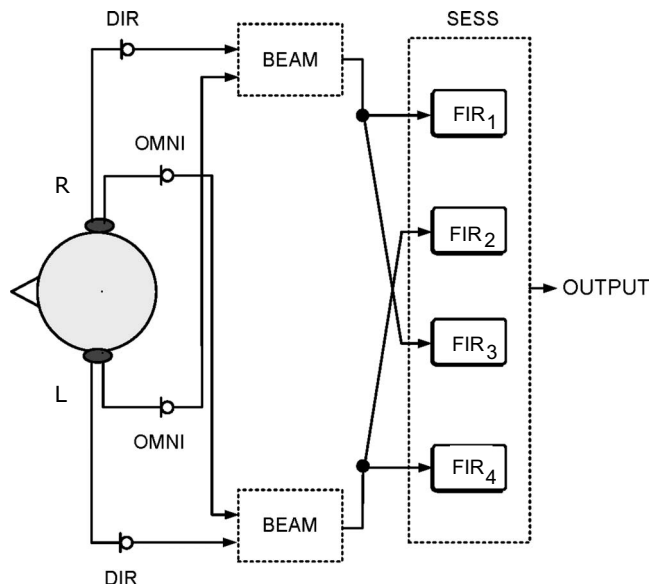

FIG. 1. Block diagram of the 4M-SESS adaptive noise reduction processing strategy.

directional (left and right) and two directional microphones (left and right) with each set of directional and omni-directional microphones placed on opposite sides of the head. Such an arrangement is already available commercially in the Nucleus Freedom implant processor.

As outlined schematically in Fig. 1, in the 4M-SESS binaural processing strategy, the speech reference on the right side is formed by adding the input to the left omni-directional microphone to the delayed version of the right directional microphone signal, and the noise reference on the right is estimated by subtracting the left omni-directional microphone signal from a delayed version of the right directional microphone signal. In a similar manner, to create the speech reference signal on the left side, the signals from the left directional microphone right and omni-directional microphone are summed together. The noise reference on the left side is formed by subtracting the right omni-directional microphone signal from a delayed version of the left directional microphone signal. Assuming that the noise source is placed on the right of the listener, this procedure leads to a signal with an amplified noise level on the right side but also yields an output with a substantially reduced noise level in the left ear. After processing the microphone signals containing two speech and two noise reference signals binaurally with one BEAM processor per ear, the two-microphone outputs from the two BEAM processors containing the generated speech and noise reference signals are fed to four adaptive linear filters and are enhanced further with the spatial enhancement via source separation (SESS) processing strategy.

The 4M-SESS strategy is the four-microphone on-line extension of the two-microphone algorithm recently proposed by Kokkinakis and Loizou (2008) and is amenable to real-time implementation. The 4M-SESS strategy operates by estimating a total of four FIR filters that can undo the mixing effect by which two composite signals are generated when the target and noise sources propagate inside a natural acoustic environment. The filters are computed after only a

K. Kokkinakis and P. C. Loizou: Coordinated stimulation in bilateral implants

TABLE I. Cochlear implant patient description and history.

	S1	S2	S3	S4	S5
Age	40	58	36	68	70
Gender	F	F	F	M	F
Years of implant experience (L/R)	6/6	4/4	5/5	7/7	8/8
Years of deafness	12	10	15	11	22
Etiology of hearing loss	Unknown	Unknown	Noise	Rubella	Hereditary

single pass with no additional training. The 4M-SESS strategy operates on the premise that the target and noise source signatures are spatially separated, and thus, their individual form can be retrieved by minimizing the statistical dependence between them (Kokkinakis and Loizou, 2010). In statistical signal processing theory, this configuration is referred to as a fully determined system, where the number of independent sound sources is equal to the number of microphones available for processing. To adaptively estimate the four unmixing filters, the 4M-SESS strategy employs the natural gradient algorithm (e.g., see Kokkinakis and Loizou, 2007, 2008). Initializing the filters used in the 4M-SESS strategy with those obtained with each of the two BEAMs, results in a substantial reduction in the total number of filter coefficients required for adequate interference rejection and speeds up the convergence of the algorithm. The implementation of the 4M-SESS strategy requires access to a single processor driving two CIs, such that signals from the left and right sides are captured synchronously and processed together.

III. METHODS

To assess the performance of the aforementioned multimicrophone noise reduction strategies, comprehensive word recognition tests were conducted with bilateral CI users.

A. Subjects

The subjects participating in the study were five American English speaking adults with postlingual deafness who received no benefit from hearing aids preoperatively. All subjects were fitted bilaterally with the Nucleus 24 multichannel implant device manufactured by Cochlear Corporation (Sydney, Australia). The participants used their devices routinely and had a minimum of 4 years experience with their CIs. Biographical data for the subjects tested is provided in Table I.

B. Research processor

All subjects tested were using the Cochlear Esprit BTE processor on a daily basis. During their visit, the participants were temporarily fitted with the SPEAR3 wearable research processor. SPEAR3 was developed by the Cooperative Research Center for Cochlear Implant and Hearing Aid Innovation (Melbourne, Australia) in collaboration with Hearworks Pty Ltd (Melbourne, Australia). The SPEAR3 has been used in a number of investigations to date as a way of controlling inputs to the cochlear implant system (e.g., see van Hoesel and Tyler, 2003). Prior to the scheduled visit of the subjects,

the Seed-Speak GUI application was used to program the SPEAR3 processor with the individual threshold (T) and comfortable loudness levels (C) for each participant.

All cochlear implant listeners used the device programmed with the advanced combination encoder speech coding strategy (Vandali et al., 2000). In addition, all parameters used (e.g., stimulation rate, number of maxima, frequency allocation table, etc.) were matched to each patient's clinical settings. The volume of the speech processor was also adjusted to a comfortable loudness prior to initial testing. Before participants were enrolled in this study institutional review board approval was obtained and before testing commenced informed consent was obtained from all participants.

C. Stimuli

The speech stimuli used for testing were sentences from the IEEE database (IEEE, 1969). Each sentence is composed of approximately seven to 12 words and in total there are 72 lists of ten sentences, each produced by a single talker. The root-mean-square amplitude of all sentences was equalized to the same root-mean-square value, which corresponds to approximately 65 dBA. The interferer was speech-shaped noise generated by approximating the average long-term spectrum of the speech to that of an adult male taken from the IEEE corpus. All the stimuli were recorded at a sampling frequency of 16 kHz.

D. Simulated reverberant conditions

A set of head-related transfer functions (HRTFs) was measured using a CORTEX MKII manikin artificial head inside a mildly reverberant room with dimensions $5.50 \times 4.50 \times 3.10$ m^3 (length $\times$ width $\times$ height) fitted with acoustical curtains to change its acoustical properties (e.g., see Van den Bogaert et al., 2009). The HRTFs were measured using identical microphones to those used in modern BTE speech processors. HRTFs provide a measure of the acoustic transfer function between a point in space and the eardrum of the listener and also include the high-frequency shadowing component due to the presence of the head and the torso. The length of the impulse responses was 4,096 sample points at 16 kHz sampling rate amounting to an overall duration of approximately 0.25 s. The reverberation time of the room was experimentally determined following the ISO 3382 standard by using an omni-directional point source B&K 4295 and an omni-directional microphone B&K 2642

TABLE II. List of spatial configurations tested.

Number of interferers	Type of interferer	Left or distributed on both sides	Right or distributed on right
One interferer	Speech-shaped noise		90°
Three interferers	Speech-shaped noise	−30°, 60°, 90°	30°, 60°, 90°

and was found to be approximately equal to 200 ms when measured in octave bands with center frequencies of 125–4000 Hz.

The artificial head was placed in the middle of a ring of 1.2 m inner diameter. Thirteen single-cone loudspeakers (FOSTEX 6301B) with a 10 cm diameter were placed every 15° in the frontal plane. A two-channel sound card (VX POCKET 440 DIGIGRAM) and DIRAC 3.1 software type 7841 (B&K Sound and Vibration Measurement Systems) were used to determine the impulse response for the left and right ear by transmitting a logarithmic frequency sweep. In order to generate the stimuli recorded at the pair of microphones for each angle of incidence the target and interferer stimuli were convolved with the set of HRTFs measured for the left- and right-hand ear, respectively. In total, there were four different sets of impulse responses for each CI configuration employing two microphones. All stimuli were presented to the listener through the auxiliary input jack of the SPEAR3 processor in a double-walled sound attenuated booth (Acoustic Systems, Inc., Houston, TX). During the practice session, the subjects were allowed to adjust the volume to reach a comfortable level in both ears.

E. Procedure

The simulated target location was always placed directly in front of the listener at 0° azimuth. Subjects were tested in conditions with either one or three interferers. In the single interferer conditions a single speech-shaped noise source was presented from the right side of the listener (+90°). In the conditions where multiple interferers were present, three interfering noise sources were placed asymmetrically either across both hemifields (−30°, 60°, and 90°) or distributed on the right side only (30°, 60°, and 90°). Table II summarizes the experimental conditions. In the single interferer condition the initial SNR was set at 0 dB, whereas for multiple interferers, the initial SNR for each interferer was fixed at 5 dB and hence, the overall level of the interferers was naturally increased as more interferers were added.

The noisy stimuli in the case of a single interferer were processed with the following stimulation strategies: (1) unilateral presentation using the unprocessed input to the directional microphone on the side ipsilateral to the noise source, (2) bilateral stimulation using the unprocessed inputs from the two directional microphones, (3) bilateral stimulation plus noise reduction using the 2M2-BEAM strategy, and (4) diotic stimulation plus noise reduction using the 4M-SESS processing strategy. The noisy stimuli generated when three interferers are originating from either the right or from both hemifields were processed with the following processing strategies: (1) bilateral stimulation using the unprocessed inputs from the two directional microphones, (2) bilateral stimulation plus noise reduction using the 2M2-BEAM strat-

egy, and (3) diotic stimulation plus noise reduction using the 4M-SESS processing strategy. In total, there were eight different conditions (4 strategies × 2 signal-to-noise-ratios) for the single interferer scenario and six different conditions (3 strategies × 2 spatial configurations) tested with multiple interferers. A total of 28 IEEE sentence lists were used. Two IEEE lists (20 sentences) were used for each condition.

Each participant completed testing in eight sessions of 2 h each spanning 2 days. None of the participants had any prior experience with the IEEE sentence material. At the start of each session, all participants were given a short practice session in order to gain familiarity with the task. Separate practice sessions were used for the single and multiple interferer conditions. No score was calculated for these practice sets. The subjects were told that they would hear sentences in a noisy background and they were instructed to type what they heard via a computer keyboard. It was explained that some of the utterances would be hard to understand and that they should make their best guess. To minimize any order effects in the test, such as learning or fatigue effects, all conditions were randomized among subjects. Different sets of sentences were used in each condition. After each test session was completed, the responses of each individual were collected, stored in a written sentence transcript and scored off-line by the percentage of the keywords correctly identified. All words were scored. The percentage correct speech recognition scores were calculated by dividing the number of key words the listener repeated correctly by the total number of key words in the particular sentence list.

IV. RESULTS AND DISCUSSION

Speech intelligibility scores obtained with the proposed multi-microphone adaptive noise reduction strategies are shown in Figs. 2 and 3 (one and three interferers, respectively).

A. One interferer

Figure 2 shows the individual subjects scores obtained using the various speech processing strategies. For comparative reasons, the scores obtained with a single (unilateral) implant are also shown. Two-way analysis of variance (ANOVA) (with repeated measures[3]) indicated significant effect [$F(2,8)=157.3$, $p<0.0005$] of the processing algorithm, significant effect of SNR level [$F(1,4)=15.9$, $p<0.05$] and significant interaction [$F(2,8)=5.2$, $p=0.035$]. The interaction is due to the fact the improvement in performance obtained with the 4M-SESS strategy, relative to the baseline bilateral condition, differed for the two SNR levels.

Post-hoc comparisons using Fisher's least significant difference (LSD) method were run to assess significant differences in scores obtained between conditions. The scores
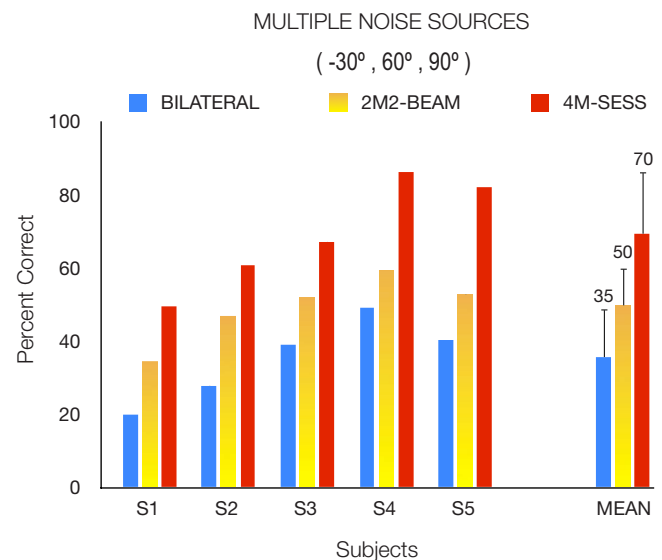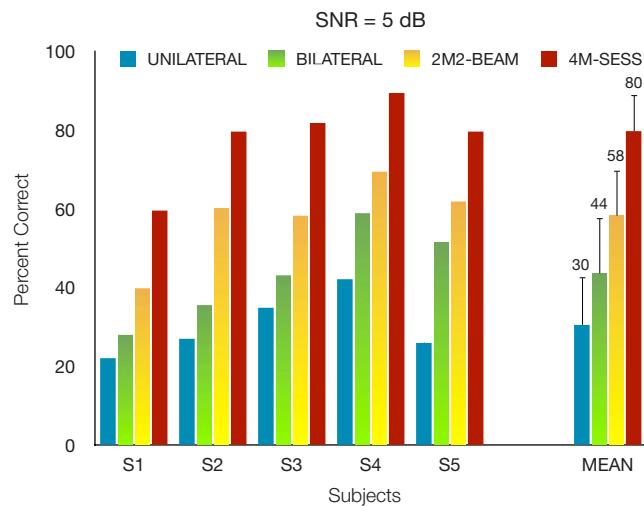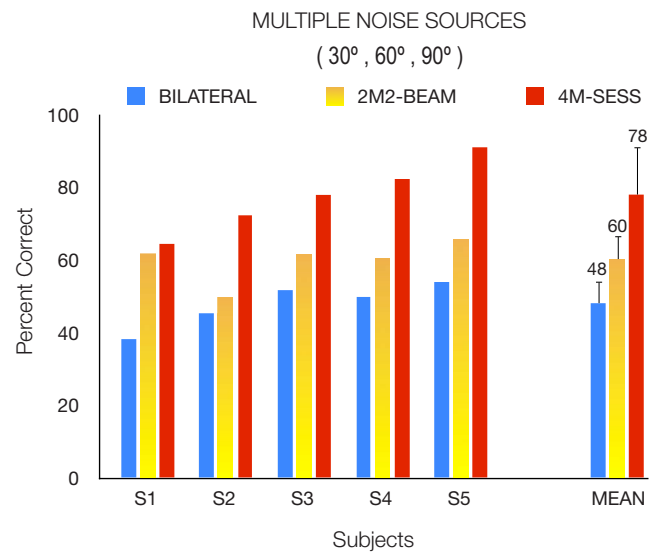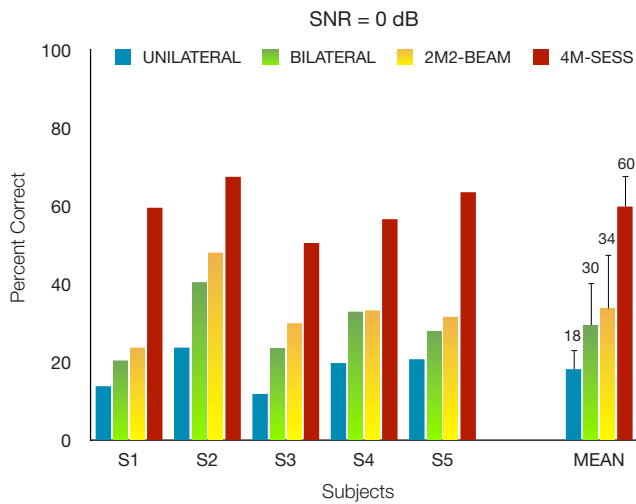
FIG. 2. (Color online) Percent correct scores by five Nucleus 24 users using both CI devices tested on IEEE sentences originating from the front of the listener (0°) and corrupted by a single speech-shaped noise source placed on the right (+90°). Error bars indicate standard deviations.

FIG. 3. (Color online) Percent correct scores by five Nucleus 24 users using both CI devices tested on IEEE sentences originating from the front of the listener (0°) and corrupted by three speech-shaped noise sources distributed on the right side only (30°, 60°, and 90°) or placed asymmetrically across both the left and right hemifields (−30°, 60°, and 90°). Error bars indicate standard deviations.

obtained with unilateral stimulation did not differ significantly ($p > 0.05$) from the scores obtained with bilateral stimulation in the 5 dB SNR condition but were significantly lower in the 0 dB SNR condition ($p = 0.026$). The scores obtained with the 2M2-BEAM strategy in the 0 dB SNR conditions did not differ significantly ($p = 0.359$) from the scores obtained with bilateral processing (baseline condition) but were significantly higher ($p = 0.036$) in the 5 dB condition. In brief, processing with the 2M2-BEAM strategy, yielded no significant benefit in intelligibility in reverberant conditions in the low SNR condition (0 dB), compared to the unprocessed bilateral condition. In contrast, the 4M-SESS processing strategy yielded a significant benefit relative to both unilateral and bilateral stimulations in both SNR conditions. More importantly, the scores obtained with the 4M-SESS strategy were significantly higher than the scores obtained with the 2M2-BEAM strategy for the 0 dB ($p < 0.005$) and 5 dB ($p = 0.004$) conditions.

B. Three interferers

Figure 3 shows the individual subjects scores obtained using the various speech processing strategies in the two

target-masker configurations. Two-way ANOVA (with repeated measures) indicated significant effect [$F(2,8) = 92.8$, $p < 0.0005$] of the processing algorithm, significant effect of interfering sources' locations [$F(1,4) = 9.24$, $p = 0.038$] and non-significant interaction [$F(2,8) = 1.18$, $p = 0.353$]. The above analysis confirms that the locations of the multiple sources (all three sources to the right or placed asymmetrically in both hemifields) had a significant effect on performance. More specifically, performance was significantly lower when the three interfering sources were placed asymmetrically in both hemifields.

Post-hoc comparisons using Fisher's LSD were run to assess significant differences in scores between the scores obtained in the (30°, 60°, and 90°) and (−30°, 60°, and 90°) spatial configurations. The scores obtained with the 2M2-BEAM processing strategy did not differ significantly (p

>0.05) from the scores obtained with bilateral (unprocessed) stimulation in the ($-30°$, $60°$, and $90°$) condition but were significantly ($p=0.026$) higher than the bilateral scores in the ($30°$, $60°$, and $90°$) condition. The scores obtained with the 4M-SESS strategy in both target-masker configurations, were significantly higher ($p<0.005$) than then scores obtained with bilateral (unprocessed) stimulation. Additionally, the scores obtained with the 4M-SESS strategy revealed a significant benefit ($p=0.003$) relative to the 2M2-BEAM in both target-masker configurations. In brief, the 2M-BEAM strategy provided no benefit (relative to the bilateral baseline condition) in intelligibility in the challenging condition where the noise sources were distributed on both sides of the user. In contrast, the 4M-SESS strategy provided consistent benefit in all conditions and in all target-masker configurations.

C. Discussion

As shown in Fig. 2, when the noise source arrives from the side ipsilateral to the unilateral implant, speech understanding in noise was quite low for all subjects when presented with the input to the fixed directional microphone. Although one would expect that processing with a directional microphone alone would result in a more substantial noise suppression this is not the case here. The small benefit observed is probably due to the high amount of reverberant energy present, since most directional microphones are designed to only work efficiently in relatively anechoic scenarios or with a moderate levels of reverberation (Hawkins and Yacullo, 1984). In contrast, improvements in speech understanding were obtained when BI-CI listeners used both cochlear implant devices. When compared to the unprocessed unilateral condition, the mean improvement observed was 12 percentage points at 0 dB SNR and 14 percentage points due to processing when the SNR was fixed at 5 dB in the single interferer scenario.

As expected, the bilateral benefit was smaller when multiple noise sources were present (see Fig. 3). There is ample evidence in the binaural hearing literature suggesting that when both devices are available, CI listeners can benefit from the so-called head-shadow effect which occurs mainly when speech and noise are spatially separated (e.g., see Shaw, 1974; Tyler *et al.*, 2003; Hawley and Litovsky, 2004; Litovsky *et al.*, 2009). For three out of five bilateral subjects tested, the scores indicated that the signal from the noise source coming from the right side was severely attenuated prior to reaching the left ear. In the case where there was only one interferer (steady-state noise) on the right ($+90°$) a better SNR was obtained at the left ear compared to the right. The individual listener was able to selectively focus on the left ear placed contralateral to the competing noise source to improve speech recognition.

Comparing the scores obtained with the 2M2-BEAM processing strategy against the baseline bilateral unprocessed condition, we observe that there is a marginal improvement in speech intelligibility. Still, the results indicate that the subjects tested were not able to significantly benefit from the use of the binaural beamformer over their daily strategy. We attribute that to the presence of reverberation. As depicted in Fig. 3, this effect was more prominent in the multiple interferer scenario. The small improvement noted can be in part attributed to the fact that the performance of beamforming techniques has been known to degrade sharply in reverberant conditions (e.g., see Greenberg and Zurek, 1992; Hamacher *et al.*, 1997; Kompis and Dillier, 1994; van Hoesel and Clark, 1995; Ricketts and Henry, 2002; Lockwood *et al.*, 2004; Spriet *et al.*, 2007). In the study by van Hoesel and Clark (1995), beamforming attenuated spatially separated noise by 20 dB in anechoic settings, but only by 3 dB in highly reverberant settings. More recently, another study by Lockwood *et al.* (2004) focusing on measuring the effects of reverberation demonstrated that the performance of beamforming algorithms is decreased by substantial amounts as the reverberation time of the room increases. Ricketts and Henry (2002) measured speech intelligibility in hearing aid users under moderate (0.3 s) and severe levels of reverberation (0.9 s) and showed that the benefit due to beamforming decreases, and in fact, almost entirely disappears near reflective surfaces or highly reverberant environments. Spriet *et al.* (2007) reached a similar outcome and concluded that reverberation had a detrimental effect on the performance of the two-microphone BEAM strategy. In accordance with the aforementioned research studies, the data from this study suggest that the binaural beamformer (2M2-BEAM) strategy cannot enhance the speech recognition ability of CI users considerably under the reverberant listening conditions tested.

In contrast, the 4M-SESS strategy yielded a considerable benefit in all conditions tested and for all five subjects. As evidenced from the scores plotted in Figs. 2 and 3, the observed benefit over the subjects' daily strategy ranged from 30 percentage points when multiple interferers were present to around 50 percentage points for the case where only a single interferer emanated from a single location in space. This improvement in performance with the proposed 4M-SESS strategy was maintained even in the challenging condition where three noise sources were present and were asymmetrically located across the two hemifields. In addition, the overall benefit after processing with the 4M-SESS processing strategy was significantly higher than the benefit received when processing with the binaural 2M2-BEAM noise reduction strategy. In both the single and multi-noise source scenarios the 4M-SESS strategy employing two contralateral microphone signals led to a substantially increased performance, especially when the speech and the noise source were spatially separated.

The observed improvement in performance with the 4M-SESS strategy can be attributed to having a more reliable reference signal, even in situations where the noise sources are asymmetrically located across the two hemifields.

Combining spatial information from all four microphones forms a better representation of the reference signal, leading to a better target segregation than that made available with only a monaural input or two binaural inputs. Having four microphones enables us to form a better estimate of the speech component and the generated noise reference. Subsequently, speech recognition performance is considerably bet-

ter. One may interpret this as introducing the better-ear advantage into the noise reduction algorithm. The implementation of the 4M-SESS processing strategy requires access to a single processor driving both cochlear implants. Signals arriving from the left and right sides need to be captured synchronously in order to be further processed together. Coordinated binaural stimulation using signals from both left and right CI devices offers a much greater potential for noise reduction than using signals from just a single CI or two independent devices. As wireless communication of audio signals is quickly becoming possible between the two sides, it may be advantageous to have the two devices in a bilateral fitting communicate with each other (Schum, 2008). An intelligent signal transmission scheme could also be designed that would enable a signal exchange between the two implant processors.

V. CONCLUSIONS

Noise reduction strategies using as many as four microphones, namely two omni-directional and two directional microphones in each of the two (one per ear) units in a coordinated stimulation mode, have not been proposed, implemented or systematically evaluated with bilateral cochlear implant users. In the present study, we assessed the benefit of coordinated bilateral stimulation mode, in which the adaptive algorithms make use of contralateral signals to form better noise references. When compared against two BEAMs (currently employed in Freedom processors) running independently in each ear, the proposed 4M-SESS multi-microphone strategy performed significantly better in all SNR conditions and in all target-masker configurations. Significant improvements were noted even in the challenging listening environment in which three interfering sources are distributed asymmetrically on both hemifields in a reverberant setting.

[1]SRT is defined as the SNR necessary to achieve 50% intelligibility. In general, high SRT values correspond to poor speech intelligibility

[2]The direct-to-reverberant ratio (DRR) is defined as the log energy ratio of the direct and reverberant portions of an impulse response and essentially measures how much of the energy arriving is due to the direct source sound and how much is due to late arriving echoes. Since, DRR depends on the source-to-listener distance, DRR decreases as perceived reverberation increases.

[3]Non-parametric statistical tests (e.g., Friedman test) were also run to assess main effects, and the same outcomes and trends in performance were observed. For instance, a significant ($p=0.007$) effect of processing algorithm on performance was noted (5 dB SNR condition) when using the non-parametric Friedman test. A similar outcome was observed in the 0 dB SNR condition. Post-hoc analysis, based on Wilcoxon's signed rank test,

indicated that the performance with the 4M-SESS processing strategy was significantly higher ($p=0.043$) than the performance with the 2M2-BEAM strategy in both SNR conditions.

Chung, K., Zeng, F.-G., and Acker, K. N. (**2006**). "Effects of directional microphone and adaptive multichannel noise reduction algorithm on cochlear implant performance," J. Acoust. Soc. Am. **120**, 2216–2227.

Ericson, H., Svard, I., Devert, G., and Ekstrom, L. (**1988**). "Contralateral routing of signals in unilateral hearing impairment: A better method of fitting," Scand. Audiol. **17**, 111–116.

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., and Haynes, D. S. (**2007**). "Horizontal plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants," Ear Hear. **28**, 524–541.

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

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., Doering, W., Mauer, G., Fleischmann, H., and Hennecke, J. (**1997**). "Evaluation of noise reduction systems for cochlear implant users in different acoustic environments," Am. J. Otol. **18**, S46–S49.

Harford, E., and Barry, J. (**1965**). "A rehabilitative approach to the problem of unilateral hearing impairment: The contralateral routing of signals (CROS)," J. Speech Hear Disord. **30**, 121–138.

Hawkins, D. B., and Yacullo, W. S. (**1984**). "Signal-to-noise ratio advantage of binau-ral hearing aids and directional microphones under different levels of reverberation," J. Speech Hear Disord. **49**, 278–286.

Hawley, M. L., and Litovsky, R. Y. (**2004**). "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," J. Acoust. Soc. Am. **115**, 833–843.

Haykin, S. (**1996**). *Adaptive Filter Theory* (Prentice Hall, New Jersey).

Hochberg, I., Boothroyd, A., Weiss, M., and Hellman, S. (**1992**). "Effects of noise and noise suppression on speech perception for cochlear implant users," Ear Hear. **13**, 263–271.

Hu, Y., Loizou, P. C., and Kasturi, K. (**2007**). "Use of a sigmoidal-shaped function for noise attenuation in cochlear implants," J. Acoust. Soc. Am. **122**, EL128–EL134.

IEEE (**1969**). "IEEE recommended practice speech quality measurements," IEEE Trans. Audio Electroacoust. **AU17**, 225–246.

Kokkinakis, K., and Loizou, P. C. (**2007**). "Signal separation by integrating adaptive beamforming with blind deconvolution," in *Independent Component Analysis and Signal Separation*, edited by M. E. Davies, C. J. James, S. Abdallah, and M. Plumbley (Springer-Berlin, Heidelberg), pp. 495–503.

Kokkinakis, K., and Loizou, P. C. (**2008**). "Using blind source separation techniques to improve speech recognition in bilateral cochlear implant patients," J. Acoust. Soc. Am. **123**, 2379–2390.

Kokkinakis, K., and Loizou, P. C. (**2010**). "*Advances in Modern Blind Signal Separation Algorithms: Theory and Applications* (Morgan & Claypool, San Rafael, CA).

Kompis, M., and Dillier, N. (**1994**). "Noise reduction for hearing aids: Combining directional microphones with an adaptive beamformer," J. Acoust. Soc. Am. **96**, 1910–1913.

Kompis, M., and Dillier, N. (**2001**). "Performance of an adaptive beamforming noise reduction scheme for hearing aid applications. II. Experimental verification of the predictions," J. Acoust. Soc. Am. **109**, 1134–1143.

Lawson, D. T., Wilson, B. S., Zerbi, M., van den Honert, C., Finley, C. C., Farmer, J. C., Jr., McElveen, J. T., Jr., and Roush, P. A. (**1998**). "Bilateral cochlear implants controlled by a single speech processor," Am. J. Otol. **19**, 758–761.

Litovsky, R. Y., Parkinson, A., and Arcaroli, J. (**2009**). "Spatial hearing and speech intelligibility in bilateral cochlear implant users," Ear Hear. **30**, 419–431.

Litovsky, R. Y., Parkinson, A., Arcaroli, J., Peters, R., Lake, J., Johnstone, P., and Yu, G. (**2004**). "Bilateral cochlear implants in adults and children," Arch. Otolaryngol. Head Neck Surg. **130**, 648–655.

Lockwood, M. E., Jones, D. L., Bilger, R. C., Lansing, C. R., O'Brien, W. D., Jr., Wheeler, B. C., and Feng, A. S. (**2004**). "Performance of time- and frequency-domain binaural beamformers based on recorded signals from real rooms," J. Acoust. Soc. Am. **115**, 379–391.

Loizou, P. C., Hu, Y., Litovsky, P., Yu, G., Peters, R., Lake, J., and Roland, P. (**2009**). "Speech recognition by bilateral cochlear implant users in a cocktail party setting," J. Acoust. Soc. Am. **125**, 372–383.

Loizou, P. C., Lobo, A., and Hu, Y. (**2005**). "Subspace algorithms for noise

reduction in cochlear implants," J. Acoust. Soc. Am. **118**, 2791–2793.

Müller, J., Schon, F., and Helms, J. (**2002**). "Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system," Ear Hear. **23**, 198–206.

Müller-Deile, J., Schmidt, B. J., and Rudert, H. (**1995**). "Effects of noise on speech discrimination in cochlear implant patients," Ann. Otol. Rhinol. Laryngol. Suppl. **166**, 303–306.

Patrick, J. F., Busby, P. A., and Gibson, P. J. (**2006**). "The development of the nucleus freedom cochlear implant system," Trends Amplif. **10**, 175–200.

Qin, M. K., and Oxenham, A. J. (**2003**). "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," J. Acoust. Soc. Am. **114**, 446–454.

Ricketts, T. A., Grantham, D. W., Ashmead, D. H., Haynes, D. S., and Labadie, R. F. (**2006**). "Speech recognition for unilateral and bilateral cochlear implant modes in the presence of uncorrelated noise sources," Ear Hear. **27**, 763–773.

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

Schum, D. J. (**2008**). "Communication between hearing aids," Adv. Audiol. **10**, 44–49.

Shaw, E. A. G. (**1974**). "Transformation of sound pressure level from free field to the eardrum in the horizontal plane," J. Acoust. Soc. Am. **56**, 1848–1861.

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

Spahr, A., and Dorman, M. (**2004**). "Performance of patients fit with advanced bionics CII and nucleus 3G cochlear implant devices," Arch. Otolaryngol. Head Neck Surg. **130**, 624–628.

Spriet, A., Van Deun, L., Eftaxiadis, K., Laneau, J., Moonen, M., Van Dijk, B., Van Wierin-gen, A., and Wouters, J. (**2007**). "Speech understanding in background noise with the two-microphone adaptive beamformer BEAM in the nucleus freedom cochlear implant system," Ear Hear. **28**, 62–72.

Stickney, G. S., Zeng, F.-G., Litovsky, R., and Assmann, P. F. (**2004**). "Cochlear implant speech recognition with speech maskers," J. Acoust. Soc. Am. **116**, 1081–1091.

Tyler, R. S., Dunn, C. C., Witt, S., and Preece, J. P. (**2003**). "Update on bilateral cochlear implantation," Curr. Opin. Otolaryngol. Head Neck Surg. **11**, 388–393.

Tyler, R. S., Gantz, B. J., Rubinstein, J. T., Wilson, B. S., Parkinson, A. J., Wolaver, A., Preece, J. P., Witt, S., and Lowder, M. W. (**2002**). "Three-month results with bilateral cochlear implants," Ear Hear. **23**, 80S–89S.

Van den Bogaert, T., Doclo, S., Wouters, J., and Moonen, M. (**2009**). "Speech enhancement with multichannel Wiener filter techniques in multimicrophone binaural hearing aids," J. Acoust. Soc. Am. **125**, 360–371.

van Hoesel, R. J. M. (**2004**). "Exploring the benefits of bilateral cochlear implants," Audiol. Neuro-Otol. **9**, 234–246.

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.

van Hoesel, R. J. M., and Clark, G. M. (**1997**). "Psychophysical studies with two binaural cochlear implant subjects," J. Acoust. Soc. Am. **102**, 495–507.

van Hoesel, R. J. M., and Tyler, R. S. (**2003**). "Speech perception, localization, and lateralization with binaural cochlear implants," J. Acoust. Soc. Am. **113**, 1617–1630.

Vandali, A. E., Whitford, L. A., Plant, K. L., and Clark, G. M. (**2000**). "Speech perception as a function of electrical stimulation rate: Using the nucleus 24 cochlear implant system," Ear Hear. **21**, 608–624.

Weiss, M. (**1993**). "Effects of noise and noise reduction processing on the operation of the nucleus 22 cochlear implant processor," J. Rehabil. Res. Dev. **30**, 117–128.

Wouters, J., and Vanden Berghe, J. (**2001**). "Speech recognition in noise for cochlear implantees with a two microphone monaural adaptive noise reduction system," Ear Hear. **22**, 420–430.

Yang, L.-P., and Fu, Q.-J. (**2005**). "Spectral subtraction-based speech enhancement for cochlear implant patients in background noise," J. Acoust. Soc. Am. **117**, 1001–1004.