informa healthcare

International Journal of Audiology 2015; Early Online: 1–9

International Journal of Audiology

Original Article

An evaluation of the performance of two binaural beamformers in complex and dynamic multitalker environments

Virginia Best[*,†], Jorge Mejia[*,‡], Katrina Freeston[*], Richard J. van Hoesel[‡,#] & Harvey Dillon[*,‡]

[*]*National Acoustic Laboratories, Australian Hearing Hub, Macquarie University, Australia,* [†]*Department of Speech, Language and Hearing Sciences, Boston University, Boston, USA,* [‡]*The HEARing Cooperative Research Centre, Carlton, Australia, and* [#]*The University of Melbourne, Parkville, Australia*

The British Society of Audiology

The International Society of Audiology

NAS NORDIC AUDIOLOGICAL SOCIETY

Abstract

Objective: Binaural beamformers are super-directional hearing aids created by combining microphone outputs from each side of the head. While they offer substantial improvements in SNR over conventional directional hearing aids, the benefits (and possible limitations) of these devices in realistic, complex listening situations have not yet been fully explored. In this study we evaluated the performance of two experimental binaural beamformers. *Design:* Testing was carried out using a horizontal loudspeaker array. Background noise was created using recorded conversations. Performance measures included speech intelligibility, localization in noise, acceptable noise level, subjective ratings, and a novel dynamic speech intelligibility measure. *Study sample:* Participants were 27 listeners with bilateral hearing loss, fitted with BTE prototypes that could be switched between conventional directional or binaural beamformer microphone modes. *Results:* Relative to the conventional directional microphones, both binaural beamformer modes were generally superior for tasks involving fixed frontal targets, but not always for situations involving dynamic target locations. ***Conclusions:*** Binaural beamformers show promise for enhancing listening in complex situations when the location of the source of interest is predictable.

Key Words: Binaural beamformers; directional hearing aids; dynamic environments

Directional amplification can increase the signal-to-noise ratio (SNR) in configurations where sounds of interest are spatially separated from interfering sounds. Recent advances in bilateral wireless technology have enabled the implementation of various highly directional binaural beamformer algorithms in hearing aids. The potential benefits from such schemes have been demonstrated in terms of acoustic SNR improvements (e.g. Greenberg & Zurek, 1992) and speech reception thresholds in noise (SRTn: Kompis & Dillier, 1994; Cornelis et al, 2012). One recent study (Picou et al, 2014) examined the benefits of binaural beamforming under conditions that better approximate real-world listening. In that study, sentence recognition for a frontal target, horizontal sound localization, and subjective ratings were measured in the presence of background noise and reverberation. The authors reported modest benefits of beamforming over conventional directional amplification for sentence recognition, poorer localization when visual information was excluded, and no effect on subjective ratings.

In the current study we evaluated the performance of two experimental beamformers, one from Phonak (Switzerland) and one from the HEARing CRC (Australia). These beamformers represent two quite different realizations, with different properties, and the aim here was not to compare them but to use them as two examples and assess their performance relative to standard directional microphones. Similar to Picou et al, we used a battery of tests measuring speech perception, spatial perception, and subjective impressions. The tests were based largely on commonly used laboratory tests, but with a few novel adaptations designed to increase the complexity of the listening scenarios. Specifically, we used competing conversations for the background noise in all tests, and considered cases in which the target of interest was not located directly in front. This allowed us to explore the performance of the beamformers under non-ideal conditions like those that might be encountered in real-world listening.

Social situations often require a listener to attend to sounds coming from a range of locations, such as when participating in a group conversation at a dinner table. Given that binaural beamformers provide a large forward looking advantage, it seems rather counter-intuitive to expect this kind of technology to preserve the same perception of surrounding sounds that people with normal hearing experience. Nonetheless, if surrounding sounds are audible and crudely locatable, listeners may be able to direct their attention and orient their heads to sounds of interest. If orienting is done optimally, then the benefit

Correspondence: Virginia Best, Department of Speech, Language and Hearing Sciences, Boston University, 635 Commonwealth Ave., Boston, MA 02215, USA. E-mail: ginbest@bu.edu

(Received 1 July 2014; accepted 2 June 2015)

ISSN 1499-2027 print/ISSN 1708-8186 online © 2015 British Society of Audiology, International Society of Audiology, and Nordic Audiological Society
DOI: 10.3109/14992027.2015.1059502

Abbreviations

4FAHL	Four-frequency average hearing loss
ANL	Acceptable noise level
SNR	Signal-to-noise ratio
SRTn	Speech reception threshold in noise

from beamforming may be maintained even when the location of the target of interest changes. In this study, speech intelligibility was measured in a relatively traditional listening situation with a fixed frontal target, as well as in a novel dynamic listening task that incorporates off-center targets (a similar test was developed recently by Jensen et al, 2012). We hypothesized that sub-optimal orienting under dynamic conditions would reduce the benefit obtained from beamforming.

An ability to locate sounds in the world is important for awareness of the environment and for navigating within it. The primary cues for spatial hearing are differences in arrival time and level of sound at the two ears (interaural time and level differences, ITD and ILD; Blauert, 1997). It is well known that these cues also provide an advantage for speech understanding in complex environments (e.g. Hawley et al, 1999). However, hearing-aid processing can alter the natural localization cues available to a listener. For example, bilateral asynchronies in directional microphones disrupt ITDs and ILDs (Keidser et al, 2006). The problem is exacerbated in many binaural beamformer algorithms, which combine multiple microphone signals to provide a single-channel output with an improved SNR. This problem can be reduced by opening the ear canal with large vents so as to increase the dominance of direct sound (containing binaural cues) transmitted to the ear canal. However, effective operation of the hearing instrument is then limited to higher frequencies, which can reduce the potential for SNR enhancement. Another option is to use a closed fitting, and employ some form of 'cue-preserving' algorithm (e.g. Desloge et al, 1997; Picou et al, 2014). In the current study, closed fittings were used for both devices to optimize SNR benefits, and one of the beamformers incorporated processing aimed at preserving some spatial information in off-center sounds. Our hypothesis was that this form of cue preservation would enable some degree of localization despite the use of a closed fitting.

While objective laboratory tests are important for estimating the benefits of hearing devices, those benefits must be perceived by the listener for the device to have a good chance of uptake and success (Dillon, 2012). In order to assess the *subjective* benefit of the two beamformers, we also incorporated a measurement of the acceptable noise level (ANL), the poorest SNR at which a person is willing to listen to speech (Nabelek et al, 1991). Recent studies suggest that the ANL tends to be positive (Olsen et al, 2012a, 2012b; Walravens et al, 2014), as opposed to SRTn measures which generally focus on negative SNRs. Thus the inclusion of the ANL test had the advantage of assessing the benefit of the beamformers at higher SNRs, where it is often argued that most real-world communication occurs (Smeds et al, 2015). Directionality has been shown to increase noise tolerance as measured by the ANL in accordance with improvements in SNR (Wu, 2010), and thus our prediction was that stronger directionality would lead to even better ANLs. We also collected subjective ratings of quality and preference, to ensure that any benefits of beamforming did not come at a cost in some other dimension that participants were sensitive to.

Methods

Participants

Twenty-seven adults with sensorineural hearing loss (19 male, 8 female) were recruited from the National Acoustic Laboratories' volunteer database. The age of the participants ranged from 30 to 79 (mean 70 years). Their hearing losses ranged from moderate to severe, and were captured for the purposes of analysis in this study by the four-frequency average hearing loss (4FAHL, mean of left- and right-ear pure-tone thresholds at 0.5, 1, 2, and 4 kHz), which ranged from 18 to 79 dB HL. All participants had relatively symmetric losses (interaural differences in the 4FAHL within 10 dB), with the exception of two participants who had asymmetries at one frequency only. Nineteen of the participants were experienced hearing-aid wearers, but none had previous experience with binaural beamformers.

Devices and fitting

Two sets of bilateral BTE hearing aids with fully occluded earmoulds were used. The first set ('Phonak') were experimental prototypes in standard casings. The second set ('CRC') used identical casings that were attached via programming cables to a laptop for real-time processing. The Phonak devices had dummy cables affixed so that the two sets of devices were indistinguishable to the participants. Both devices included independently operating cardioid frontal-looking directional microphones as a reference condition, whose directivity patterns were matched as closely as possible across the devices, and were programmed with their respective experimental binaural beamformers.

The Phonak beamformer combines the four omnidirectional microphones from the left and right hearing aids with the following binaural processing scheme. First, on each side, the two microphones are processed to obtain a standard front facing cardioid-type beam. Then these directional signals are exchanged over the wireless link with the other hearing aid. Utilizing a frequency-dependent weighting function, each hearing aid then linearly combines the ipsilateral and contralateral directional signals to create a binaural directivity. The binaural beamwidth is controlled by the weighting function and is typically narrower than what a simple monaural two-microphone beamformer is able to achieve. The final step is the extension of this static binaural beamformer scheme to the adaptive binaural beamformer that optimally adapts the binaural directivity to the present spatiotemporal distribution of noise sources. This is accomplished by a well-known generalized sidelobe structure (Griffiths & Jim, 1982) which is used to adaptively combine the static binaural beamformer output with a directional signal calculated from the ipsilateral and contralateral microphone signals.

The CRC beamformer was one variation of an ongoing development at the HEARing CRC, Australia. Earlier variants of the beamformer are described in Mejia et al (2009) and van Hoesel and Mejia (2011). In all these approaches left and right microphone signals are evaluated to determine the extent to which signals from the microphones at the two ears differ. The measure of difference is used to design an adaptive filter that reduces the estimated contribution from signals that originate outside the median plane. The strength of that filter can be adjusted, with the trade-off that more aggressive filtering increases the likelihood of target distortion, particularly at low SNRs. The beamwidth of the filter is generally narrowed with more aggressive filtering, but additional parameters are available to compensate for that effect. Because the filter is adaptive, the beamwidth cannot be specified for arbitrary signals and SNRs. The beamwidth of the variation tested here was therefore subjectively adjusted using several sound fields so that for positive SNRs signals appeared well retained if they arrived from within ± 20° of the look direction, and were clearly attenuated beyond about ± 45°. However with increasingly negative SNRs target signals are more often also affected by filtering even for targets directly to the front. Accordingly, the beamwidth and filter strength settings were also constrained by the requirement to keep those occurrences to an acceptable level for SNRs in the range − 5 to − 10 dB. Because this version of the CRC beamformer produced a single channel output, a further 'cue-preserving' step was included in which an attenuated version of the original unprocessed binaural signal was mixed in with the processed signal in an attempt to retain some of the spatial information contained in the sound field.

At the first appointment otoscopy and pure-tone audiometry were performed and impressions were taken to allow for custom earmoulds to be fitted. At the beginning of the second appointment both devices were fitted to the participant's earmoulds. Insertion gain measures were performed in the directional program, using speech-shaped noise, and adjustments were made to match NAL-RP targets from 0.25 to 6 kHz. These (linear) gain settings were then copied into the beamformer program. Where the NAL-RP prescription did not prescribe gain in the low frequencies, 5 dB of amplification was provided to reduce possible effects of direct sound leakage. All other adaptive features (e.g. noise reduction) were deactivated.

Environment

Testing was carried out in a large test booth ($T_{60} = 0.3$ s) containing a horizontal array of 16 loudspeakers (Genelec 8020C) arranged in a circle with a radius of 1.2 m. The loudspeakers were driven by two ADI-8 DS digital-to-analogue converters and an RME Fireface UFX interface (44.1 kHz output). The participant was seated at the center of the array in a chair, with their head at approximately the height of the loudspeakers. A speech background was created using scripted conversations recorded previously under anechoic conditions. Each conversation was between two different talkers, which could be both male, both female, or mixed gender. In most conditions, seven different conversations were presented at seven loudspeaker locations (± 45°, ± 90°, ± 135°, and 180° azimuth). The two talkers in each conversation were presented from the same loudspeaker. At any moment there were seven simultaneous talkers, each normalized to the same root-mean-square level, with a combined total level of 65 dB SPL (measured in the center of the array with no participant present).

Test battery

Testing was done over 3–4 visits of approximately two hours each. Where possible, all participants completed the test battery under each of the four hearing-aid conditions. The order of testing of the four hearing-aid conditions was counterbalanced across participants for all tests. No systematic training was included, but clear verbal instructions were given before commencing each test.

Speech intelligibility in noise

Speech-in-noise testing was done using a large corpus of sentences based on the Bamford-Kowal-Bench sentences (BKB; Bench & Bamford, 1979). The sentences were spoken by a single male talker with Australian accented English, and were presented from the loudspeaker directly in front (0° azimuth). Participants were instructed to keep their head still and facing this loudspeaker, and compliance was monitored from the observation area via a video camera. The task was to repeat each sentence aloud into a microphone. Responses were conveyed to the investigator, who logged the correct morphemes using custom-made software.

In one condition, background noise consisted of the 7-talker background described above. In a second condition, a sparser 3-talker background was used in which three conversations were located at ± 45° and 90° azimuth, again with an overall level of 65 dB SPL. In a third condition, a symmetric 2-talker background was used with two conversations at ± 90° azimuth, again with an overall level of 65 dB SPL. This last condition was only completed by a subset of eight participants who were willing to attend an additional testing session.

For each background noise environment, one adaptive track (see Keidser et al, 2013b) was initially completed in the Phonak directional mode to estimate the target level at which 50% correct identification was achieved. This SRTn was used to select an SNR for fixed-level testing. Pilot testing revealed that the SNR needed to be lowered by about 2 dB to place fixed-level performance at 50% correct, either because of a practice effect or because of differences between adaptive and fixed-level testing, and thus SRTn-2dB was selected for each listener. Across listeners this SNR ranged from − 10 to + 1 dB in the 7-talker background, from − 12 to + 1 dB in the 3-talker background, and from − 15 to − 2 dB in the 2-talker background. Two 32-trial blocks were obtained at the chosen SNR for each of the four hearing-aid conditions. Performance in percent correct was calculated for each of these two blocks and averaged.

Dynamic speech perception

This test followed a similar format to that described above for speech intelligibility in the 7-talker environment. Again a fixed SNR was chosen for each participant, but it was set at SRTn (i.e. 2 dB above that used in the previous test), with the expectation that performance would be poorer overall in the dynamic task. In this task the first target sentence was presented from the frontal loudspeaker (0° azimuth), the second sentence randomly from one of four off-center locations (± 22.5° and ± 67.5° azimuth), the third sentence from the frontal loudspeaker, etc. An LED above the appropriate loudspeaker was illuminated half a second prior to each target sentence to alert participants to the new location, and they were instructed to turn their head to maximize speech reception. The experimenter ensured that participants were making head movements via the video camera. Two 32-trial blocks were obtained for each of the four hearing-aid conditions. Two participants did not have time to complete the dynamic task.

Localization in noise

Localization of a 200-ms white noise burst was examined in the 7-talker background. The noise was high-passed at 0.5 kHz (upper cut-off: 22.05 kHz) to minimize potential effects related to low-frequency sounds entering the ear canal via a direct path. The target was presented randomly from the 16 loudspeakers, which were labelled with the numbers 1 − 16 according to a schematic that was given to the participant prior to testing. A target level of 65 dB SPL was chosen such that it was audible for the majority of participants, but for three participants it was clear that this level was completely inaudible in at least one hearing-aid condition, and thus the level of the target was increased by up to 5 dB as needed. One participant still had trouble consistently hearing the target at 70 dB and thus results are not included for this participant.

Participants were instructed to keep their head still and facing the frontal loudspeaker until they had heard the target, then they were free to turn their head to read the labels or to look down at the schematic. They responded verbally with the number of the loudspeaker they judged to be the source of the sound, and were instructed to guess if they were uncertain. The experimenter recorded the response via custom software. Five responses were collected at each loudspeaker location, for a total of 80 trials per hearing-aid condition.

Acceptable noise level

The ANL test was based on previous studies (e.g. Walravens et al, 2014). Participants were given a handheld wireless keypad that was programmed to change the level of stimuli presented to the loudspeaker array. First, a monologue spoken by the same male voice used for intelligibility testing was presented from the frontal loudspeaker (in quiet). The level was initially set to 40 dB SPL, and participants were asked to adjust the speech level by first increasing the level beyond what they would like, and then decreasing it to below their preferred level, and then increasing it again to find their most comfortable level (MCL). The 7-talker background was then introduced at 40 dB SPL, and participants were instructed to increase the level of the background until they could not follow the target, decrease the level until the target was easy to follow, and then increase it again until they found the most noise they could tolerate while still following the monologue for a long period of time without becoming tense or tired (background noise level, BNL). The procedure was performed twice for each hearing-aid condition, and the MCL and BNL values were averaged. The acceptable noise level (ANL) was defined as MCL − BNL. One participant did not have time to complete the ANL task.

Subjective ratings

For each hearing-aid condition, the same monologue used above was presented from the frontal loudspeaker in the 7-talker background at a fixed SNR of 0 dB. Participants were asked to make ratings about their perception of the background noise, the target clarity, and the overall sound quality, using numerical scales (see Appendix 1). Because it was not possible to switch between the Phonak and CRC devices in real time, a set of ratings was collected for the directional and beamformer programs in one device, as well as an A-B comparison assessing overall preference, before moving to the other device. The order of testing of the two devices was counterbalanced across participants. One participant did not have time to complete the subjective ratings.

Results

Speech intelligibility in noise

The top row of Figure 1 shows average scores in percent correct for the four hearing-aid conditions in the different backgrounds. The bottom row shows the 'beamformer benefit' in each case, which is simply the difference between scores in the beamformer conditions and their respective directional conditions. In the 7-talker background (top left), scores in the directional conditions were close to 50%, and scores in the beamformer conditions were slightly higher. The beamformer benefit (bottom left) was around 7 percentage points for both devices. Paired t-tests indicated that the benefits were significant for both devices [Phonak: $t(26) = 3.45$, $p = 0.002$; CRC: $t(26) = 3.40$, $p = 0.002$]. In the 3-talker background (top middle), scores in the directional conditions were again close to 50%, and scores for the beamformer condition were somewhat higher. The beamformer benefit (bottom middle) was around 6 and 11 percentage points in the Phonak and CRC devices, respectively. The benefits were again significant for both devices [Phonak: $t(26) = 3.23$, $p = 0.003$; CRC: $t(26) = 4.92$, $p < 0.001$]. For the subset of eight participants who completed testing with the 2-talker background, scores in the directional conditions were not equivalent (top right); the score for the Phonak device was 12 percentage points lower than for the CRC

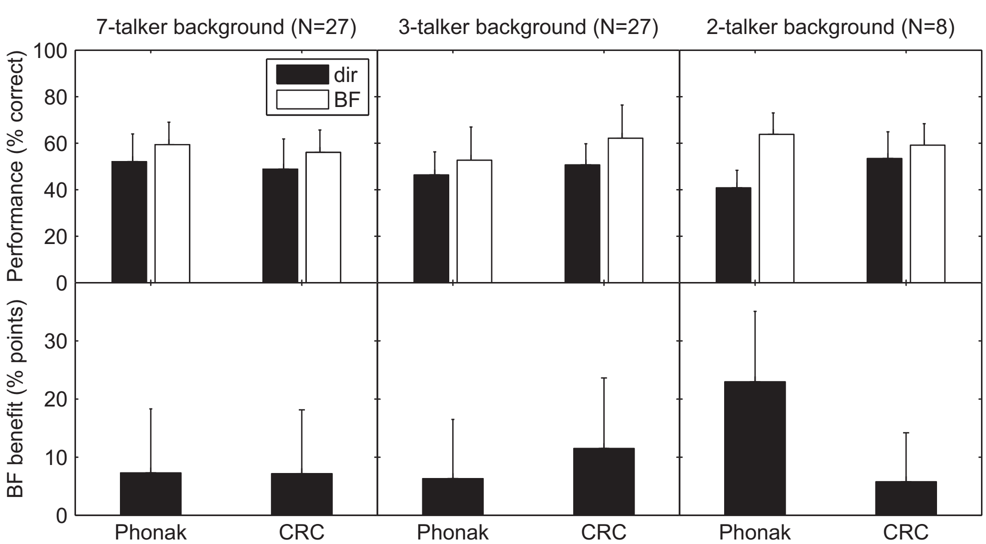

Figure 1. Top row: Group mean speech intelligibility performance for the four hearing-aid conditions in the 7-talker background, the 3-talker background, and the 2-talker background. Bottom row: Group mean beamformer benefits for the two devices in the 7-talker background, the 3-talker background, and the 2-talker background. Error bars show across-subject standard deviations.

Table 1. Simple correlations between SRT benefit and each of age, 4FAHL, and SNR at which the test was conducted.

	7-talker background		*3-talker background*	
	Phonak	*CRC*	*Phonak*	*CRC*
Age	$r = -0.26, p = 0.20$	$r = 0.12, p = 0.57$	$r = -0.34, p = 0.09$	$r = -0.04, p = 0.83$
4FAHL	$r = 0.37, p = 0.06$	$r = -0.16, p = 0.42$	$r = 0.48, p = 0.01$	$r = 0.49, p = 0.009$
SNR	$r = 0.22, p = 0.28$	$r = -0.23, p = 0.25$	$r = 0.40, p = 0.04$	$r = 0.55, p = 0.003$

device. On the other hand, scores in the two beamformer conditions were similar. The beamformer benefit (bottom right), relative to the individual directional reference, was around 23 and 6 percentage points in the Phonak and CRC devices, respectively. The benefit was significant for the Phonak device [$t(7) = 5.37$, $p = 0.001$] but not the CRC device [$t(7) = 1.93$, $p = 0.095$].

An inspection of individual data revealed a large range of binaural benefits. To explore this further, benefits in the two backgrounds for which all subjects participated were examined more closely. In the 7-talker background, benefits ranged from −12 to 30 percentage points across listeners and devices, and this range was even larger in the 3-talker background (−27 to 31 percentage points). While the correlation between individual performance benefit on the two different noise backgrounds did not reach significance [Phonak: $r = 0.36$, $p = 0.07$; CRC: $r = 0.38$, $p = 0.05$], the correlation between individual performance benefit across the two devices did [7-talker background: $r = 0.48$, $p = 0.01$; 3-talker background: $r = 0.49$, $p = 0.01$].

Analyses were conducted to examine whether the measured beamformer benefits were related to hearing loss (4FAHL), age, or the individualized SNR used for testing. Simple correlations revealed that both hearing loss and SNR were positively correlated with benefits for both devices in the 3-talker background (Table 1). However, multiple regression analyses did not find any of the variables to be significant predictors of benefit for either device in either background condition (Table 2).

Dynamic speech perception

Figure 2 shows performance on the dynamic speech task as a function of target location. For the frontal target location, it was expected that the general pattern of performance would resemble that seen in the speech intelligibility test. Indeed, performance for the directional conditions was 73% on average (above 50% as expected since the SNR was 2 dB more favourable for this test), and performance for the beamformer conditions was slightly better (77% on average). Performance worsened for all hearing-aid conditions for off-center target locations, especially at ±67.5°. This drop was most marked for CRC beamformer, which showed a clear disadvantage for these more lateral targets. A repeated measures ANOVA confirmed that the effect of device was significant [$F(3,72) = 16.5$, $p < 0.001$], as was the effect of location [$F(12,288) = 14.8$, $p < 0.001$], and the interaction between the two [$F(3,72) = 16.5$, $p < 0.001$]. Planned comparisons ($p < 0.05$) indicated that the beamformers outperformed their directional references at 0°, but underperformed at −67.5° (Phonak and CRC) and +67.5° (CRC only). Differences at ±22.5° were not significant.

Localization in noise

Two performance measures were extracted from the raw localization data. The first was a simple count of the number of front-back errors, in which responses fell in the front-back hemifield opposite to the actual target location. Front-back error rates (Figure 3, top left) were 23% and 25% for the two directional conditions, and increased to 25% and 37% for the Phonak and CRC beamformers, respectively. The increase in error (bottom left) was significant for the CRC device [$t(25) = 6.15$, $p < 0.001$] but not the Phonak device [$t(25) = 1.81$, $p = 0.082$] and was significantly larger for the CRC device [$t(25) = 3.47$, $p = 0.002$]. While these front-back error rates are relatively high, this is not unexpected given that the hearing aids have a limited bandwidth, and that these devices were BTEs, both of which serve to reduce the usefulness of pinna-related cues for front-back discrimination (Best et al, 2010).

Of more interest in the present study were errors in the left-right dimension, defined as the root-mean-square of the differences between the target and response lateral angles (regardless of front-back position). On average these errors were 32° and 30° in the directional conditions (top right), and increased substantially for both the Phonak and CRC beamformers (to 52° and 41°, respectively). The increase in error (bottom right) was significant for both devices [Phonak: $t(25) = 11.69$, $p < 0.001$; CRC: $t(25) = 8.43$, $p < 0.001$] and was significantly larger for the Phonak device [$t(25) = 3.00$, $p = 0.001$]. To examine these errors in more detail, Figure 4 shows the left-right errors as a function of lateral angle. Here it can be seen that the increased error for the beamformers comes from the more lateral locations (beyond 45°), where responses tended to be drawn towards the center. This effect is less dramatic for the

Table 2. Results of multiple regression analyses on SRT benefit data using age, 4FAHL, and SNR as predictors.

	7-talker background		*3-talker background*	
	Phonak	*CRC*	*Phonak*	*CRC*
Age	$ß = -0.11, p = 0.91$	$ß = 0.20, p = 0.40$	$ß = -0.23, p = 0.25$	$ß = 0.05, p = 0.81$
4FAHL	$ß = 0.49, p = 0.23$	$ß = 0.27, p = 0.53$	$ß = 0.31, p = 0.38$	$ß = 0.16, p = 0.64$
SNR	$ß = -0.19, p = 0.63$	$ß = -0.45, p = 0.27$	$ß = 0.11, p = 0.74$	$ß = 0.42, p = 0.22$
Overall model fit	$R^2_{adj} = 0.06, p = 0.23$	$R^2_{adj} = -0.04, p = 0.57$	$R^2_{adj} = 0.18, p = 0.06$	$R^2_{adj} = 0.21, p = 0.04$

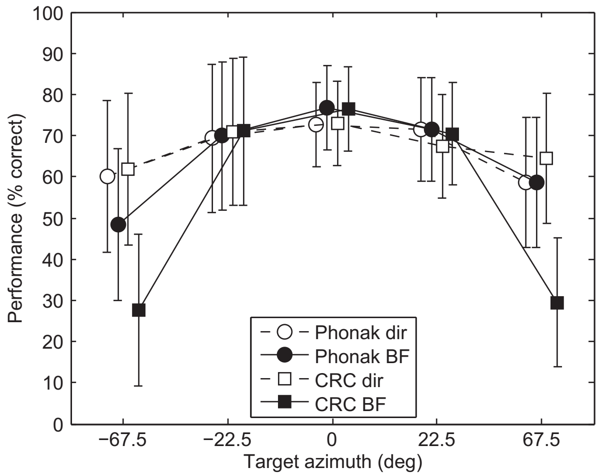

Figure 2. Group mean performance as a function of target azimuth in the dynamic speech intelligibility test for the four hearing-aid conditions. Error bars show across-subject standard deviations.

cue-preserving CRC beamformer, despite its strong attenuation of these sources.

Acceptable noise level

Mean ANLs are shown in Figure 5 (top) for the four hearing-aid conditions. For the two directional conditions the ANLs were 1.7 and − 0.5 dB, significantly lower than has been reported in other studies (i.e. more noise is tolerated), perhaps because of the noise-reducing power of directional microphones and/or the specifics of our test stimuli. The Phonak and CRC beamformers resulted in ANLs of − 0.5 and − 4.1 dB (2.2 and 3.6 dB lower than their directional

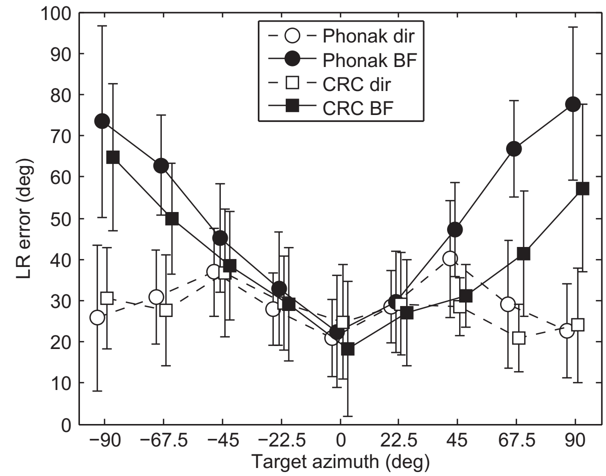

Figure 4. Group mean left-right localization errors errors as a function of target location for the four hearing-aid conditions. Error bars show across-subject standard deviations.

counterparts). These reductions (bottom) were both significant [Phonak: $t(25) = 4.15$, $p < 0.001$; CRC: $t(25) = 9.50$, $p < 0.001$].

Subjective ratings

The top row of Figure 6 shows mean subjective ratings for each device (where higher numbers correspond to more positive ratings) describing the listener's perception of the background noise, the target clarity, and the overall sound quality. The bottom row shows the difference between ratings in the beamformer conditions and their respective directional conditions. Ratings were generally similar for the two directional conditions, and generally increased for the

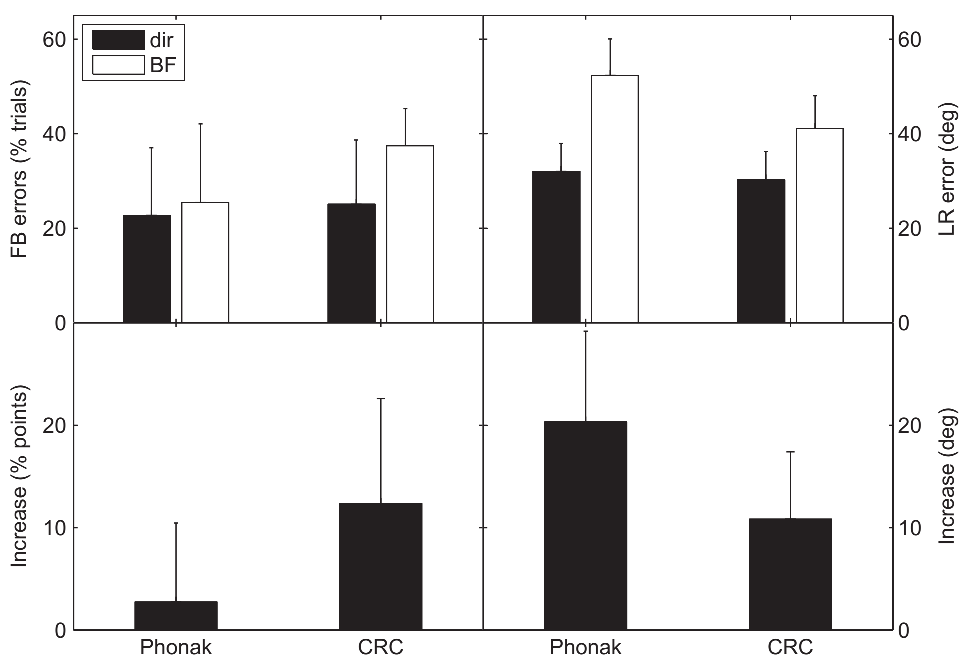

Figure 3. Top row: Group mean front-back error rates and left-right errors for the four hearing-aid conditions. Bottom row: Group mean increases in front-back error rates and left-right localization errors in the beamformer mode relative to the directional mode for the two devices. Error bars show across-subject standard deviations.

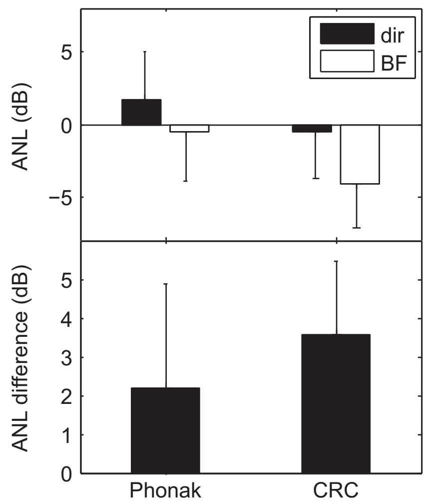

Figure 5. Top: Group mean ANLs for the four hearing-aid conditions. Bottom: Group mean decreases in ANL in the beamformer mode relative to the directional mode for the two devices. Error bars show across-subject standard deviations.

beamformer conditions. Wilcoxon signed rank tests revealed that these increases were significant for both devices for background noise [Phonak: $p = 0.01$; CRC: $p < 0.001$] and target clarity [Phonak: $p = 0.02$; CRC: $p < 0.001$] and that there were no significant differences for overall sound quality [Phonak: $p = 0.37$; CRC: $p = 0.06$].

For each listener and each set of devices, a preference rating was obtained where -3 indicates a strong preference for the directional mode, $+3$ indicates a strong preference for the beamformer mode, and 0 represents no preference. Of the 26 listeners, 16 showed a preference for the beamformer (i.e. preference rating > 0) when wearing the Phonak devices, and 22 showed a preference for the beamformer when wearing the CRC devices. The average magnitude of the preference rating was $+0.8$ and $+1.8$ for the Phonak and CRC devices, respectively. According to the rating scale (Appendix A), this indicates that as a group listeners found the CRC beamformer to be 'better' than its directional reference, and the Phonak beamformer to be 'slightly better' than its directional reference.

Discussion

In the present study, we evaluated the performance of two experimental binaural beamformers in multitalker backgrounds, using a range of objective and subjective performance measures. Both binaural beamformers were superior on average to the conventional directional microphones on many, but not all, of the performance measures.

For speech intelligibility in noise, the beamformers generally outperformed the directional microphones. In the 7-talker background both beamformers provided modest benefits. Moreover individual analysis showed that not all participants received a benefit (in fact in a few cases performance got worse). There was also a large range of benefits in the 3-talker and 2-talker conditions. Random measurement error may account for some of the individual variations in benefit as shown by Keidser et al (2013a). It is also possible that distortions of the target speech introduced by the beamforming reduced the potential benefit, more so for some listeners than others. Distortions are more likely at negative SNRs, and thus one might expect that the listeners who were tested at lower SNRs might obtain the least benefit. The correlations shown in Table 1 are consistent with this, but for the 3-talker background only, and the results of the multiple regression analysis (Table 2) suggests that the SNR would not be able to reliably predict benefits. The effect of SNR warrants further investigation, especially as the power of the regression analysis may have been affected by multicollinearity and by the relatively small sample size.

Our speech intelligibility results are broadly consistent with another recent study that measured speech intelligibility for a binaural beamformer (Picou et al, 2014). In that study, performance with binaural beamforming outperformed that with conventional directional processing for a set-up with four sources of competing multitalker babble, but only in a moderately reverberant room ($T_{60} = 0.65$ s, more reverberant than our test booth) and not in a

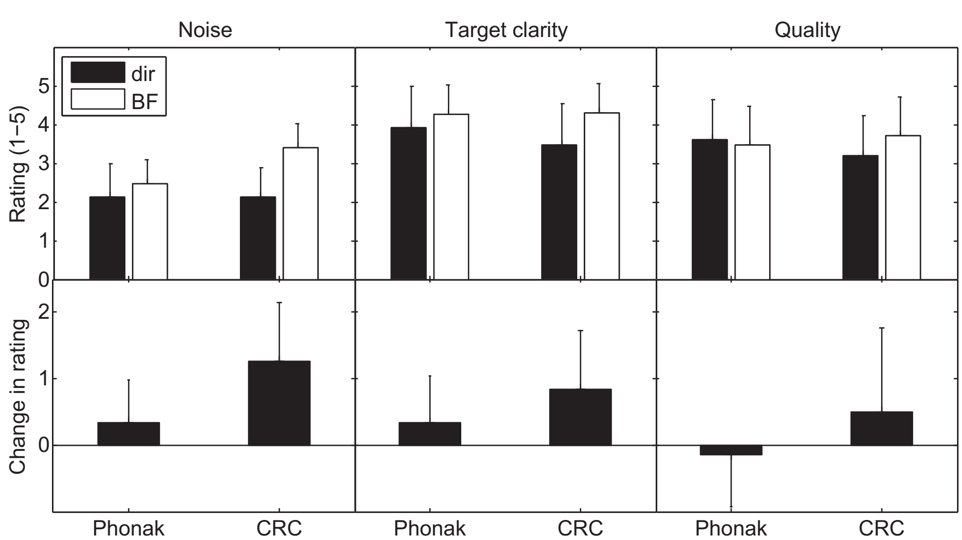

Figure 6. Top row: Group mean subjective ratings for the four hearing-aid conditions in the categories of perception of noise, target clarity, and sound quality. Bottom row: Group mean changes in subjective ratings in the beamformer mode relative to the directional mode for the two devices in the different categories. Error bars show across-subject standard deviations.

mildly reverberant room (T_{60} < 0.1 s, less reverberant than our test booth). Moreover, listeners were tested at two different SNRs, and there was a trend for the benefit to be larger at the higher SNRs (12 vs. 6 rationalized arcsine units on average). Taken together, the results suggest that binaural beamformers can provide a speech intelligibility benefit in multitalker environments, but the magnitude of this benefit depends on the specific beamformer implementation, the arrangement of competing sound sources, the reverberation present, the SNR, and the individual listener. In general, the substantial individual differences in benefit, and our inability to identify clear predictors of this benefit, are reminiscent of previous studies that examined the benefit of directional microphones (e.g. Ricketts & Mueller, 2000; Keidser et al, 2013a).

To our knowledge, very few previous studies have attempted to measure speech intelligibility under more dynamic conditions (but see Jensen et al, 2012), and none have examined the effect of directionality. Our results suggest that directionality reduces performance for lateral sources. One explanation for this relates to the arrangement of the seven masker sources in our experimental set up. For the frontal target, the adjacent maskers were separated from the target by 45°, whereas for the most lateral target, the adjacent maskers were separated from the target by 22.5°. For the intermediate target, the adjacent maskers were asymmetrically separated from the target by 22.5° and 67.5°. In other words, the inferior performance we observed for the most lateral targets may be due to very close maskers in that case, which may be harder to exclude from the directional beam (although the narrowly tuned CRC beamformer should cope relatively well with this, and yet it showed the poorest performance). Another possible explanation for the pattern of results is that listeners did not move their heads optimally when turning to face the most lateral targets, leaving the target source slightly outside the ‘sweet spot’ of the beamformer. It would be necessary to include head-tracking in future implementations of this task to determine whether the reduction in performance can be explained on the basis of imperfect (or insufficiently rapid) head orientation to the lateral sources.

The primary goal of the localization in noise task was to examine whether the CRC beamformer, which is designed to preserve some spatial information for off-center sounds, enabled some degree of localization in the left-right dimension. Localization was poor overall for both beamformers, and worse than for the conventional directional conditions, which is consistent with the results of Picou et al (2014). Also consistent with that study, the error primarily increased for the more lateral locations (beyond 45°). However, the increase in lateral errors was significantly smaller for the cue-preserving CRC beamformer. Thus for closed fittings, these data confirm that there is good reason to further investigate the benefit of preserving spatial cues under realistic listening conditions, and to explore any trade-offs with SNR (and speech intelligibility).

We included several subjective measures in our test battery to determine whether listeners actually perceived any benefits corresponding to those measured objectively. The improvement in noise acceptance observed in the ANL task for both beamformers is encouraging because this task operates at slightly higher and more realistic SNRs than the speech intelligibility task. Moreover, the subjective ratings suggested that both beamformers tended to improve subjective impressions of noise and target clarity, and did not worsen overall quality.

Finally, when deciding whether a new hearing-aid scheme offers a clear improvement over an existing one, it is often useful to consider two criteria. One is whether the new scheme leads to an improvement in some objective measure of performance, and the second is whether

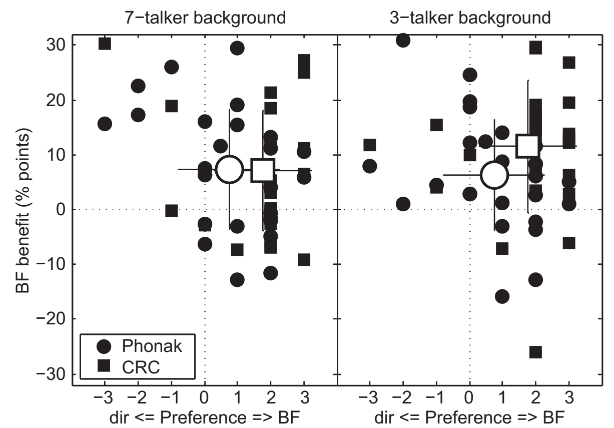

Figure 7. Individual beamformer benefits as a function of preference rating for the two devices in the 7-talker background (left) and the 3-talker background (right). The large symbols and error bars show across-subject means and standard deviations.

the new device is preferred by listeners. Support for the beamformers tested in the current study might thus come jointly from a positive improvement in speech intelligibility (relative to conventional directional microphones), and a positive preference rating. More conservatively, support might come from a benefit in either the subjective or the objective dimension, and no deficit in the other. Scatter plots of these two measures are shown for the 7-talker and 3-talker backgrounds in Figure 7, with means shown by the large open symbols. In this representation, improvements in both measures leads to data in the top right-hand quadrant, declines in both measures leads to data in the bottom left-hand quadrant, with the center of the plot indicating no change in performance and ‘neutral’ preference. For both devices, under both background conditions, the majority of points (and the means) fall in the top-right quadrant, suggesting that overall the beamformers do offer an improvement over conventional directional microphones. We currently do not have a good explanation for the exceptions, such as listeners with a clear preference that does not match their intelligibility benefit. It is also worth noting that intelligibility benefits and positive preference ratings collected in the laboratory do not always translate into positive preferences in real-world listening situations (e.g. Cord et al, 2004, 2007; Picou et al, 2014).

Conclusion

In this study we evaluated the performance of two binaural beamformers, relative to conventional directional hearing aids, for several objective and subjective listening tasks in multitalker backgrounds. Both beamformers offered improvements in performance for tasks involving fixed, frontal targets, but performance was generally not improved (and was at times disrupted) under more dynamic conditions.

Acknowledgements

The authors would like to thank Markus Hofbauer and Gabriel Koenig for their critical input, and Cong Van Nguyen for technical assistance.

Declaration of interest: This work was supported by Phonak AG and the Australian Government Department of Health. Virginia Best was also partially supported NIH/NIDCD grants DC04545 and DC013286. Portions of the work were presented at the World Congress of Audiology, Brisbane, Australia, May 2014.

References

Bench J. & Bamford J.M. (eds.) 1979. *Speech-hearing Tests and the Spoken Language of Hearing-impaired Children*. London and New York: Academic Press.

Best V., Kalluri S., McLachlan S., Valentine S., Edwards B. et al. 2010. A comparison of CIC and BTE hearing aids for three-dimensional localization of speech. *Int J Audiol*, 49, 723–732.

Blauert J. 1997. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge: MIT Press.

Cord M., Baskent D., Kalluri S. & Moore B.C.J. 2007. Disparity between clinical assessment and real-world performance of hearing aids. *Hear Rev*, 14, 22–26.

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

Cornelis B., Moonen M. & Wouters J. 2012. Speech intelligibility improvements with hearing aids using bilateral and binaural adaptive multichannel Wiener filtering based noise reduction. *J Acoust Soc Am*, 131, 4743–4755.

Desloge J.G., Rabinowitz W.M. & Zurek P.M. 1997. Microphone-array hearing aids with binaural output. I. Fixed-processing systems. *IEEE Trans. Speech Audio Process*, 5, 529–542.

Dillon H. 2012. *Hearing Aids*. Turramurra: Boomerang Press.

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

Griffiths L. & Jim C. 1982. An alternative approach to linearly constrained adaptive beamforming. *IEEE Trans. on Antennas and Propagation*, 30, 27–34.

Hawley M.L., Litovsky R.Y. & Colburn H.S. 1999. Speech intelligibility and localization in a multi-source environment. *J Acoust Soc Am*, 105, 3436–3448.

Jensen N.S., Johannesson R.B., Laugesen S. & Hietkamp R.K. 2012. Measuring speech-in-speech intelligibility with target location uncertainty. In: T. Dau et al. (ed.) *Speech perception and auditory disorders. Proceedings of the International Symposium on Audiological and Auditory Research (ISAAR)*, pp. 135–142.

Keidser G., Dillon H., Convery E. & Mejia J. 2013a. Factors influencing inter-individual variation in perceptual directional microphone benefit. *J Am Acad Audiol*, 24, 955–968.

Keidser G., Dillon H., Mejia J. & Nguyen C.V. 2013b. An algorithm that administers adaptive speech-in-noise testing to a specified reliability at selectable points on the psychometric function. *Int J Audiol*, 52, 795–800.

Keidser G., Rohrseitz K., Dillon H., Hamacher V., Carter L. et al. 2006. The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers. *Int J Audiol*, 45, 563–579.

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

Mejia J., Carlile S. & Dillon H. 2009. Method and system for enhancing the intelligibility of sounds. US 2009/0304188 A1.

Nabelek A.K., Tucker F.M. & Letowski T.R. 1991. Toleration of background noises: Relationship with patterns of hearing-aid use by elderly persons. *J Sp Hear Res*, 34, 679–685.

Olsen S.Ø., Lantz J., Nielsen L.H. & Brännström K.J. 2012a. Acceptable noise level (ANL) with Danish and non-semantic speech materials in adult hearing-aid users. *Int J Audiol*, 51, 678–688.

Olsen S.Ø., Nielsen L.H., Lantz J. & Brännström K.J. 2012b. Acceptable noise level: Repeatability with Danish and non-semantic speech materials for adults with normal hearing. *Int J Audiol*, 51, 146–156.

Picou E.M., Aspell E. & Ricketts T.A. 2014. Potential benefits and limitations of three types of directional processing in hearing aids. *Ear Hear*, 35, 339–352.

Ricketts T. & Mueller H.G. 2000. Prediciting directional hearing-aid benefit for individual listeners. *J Am Acad Audiol*, 11, 561–569.

Smeds K., Wolters F. & Rung M. 2015. Estimation of signal-to-noise ratios in realistic sound scenarios. *J Am Acad Audiol*, 26, 183–196.

Van Hoesel R. & Mejia J. 2011. Systems and methods for reducing unwanted sounds in signals received from an arrangement of microphones. PCT/AU2011/001476.

Walravens E., Keidser G., Hartley D. & Hickson L. 2014. An Australian version of the acceptable noise level test and its predictive value for successful hearing-aid use in an older population. *Int J Audiol*, 53, S52–59.

Wu T.-H. 2010. Effect of age on directional microphone hearing-aid benefit and preference. *J Am Acad Audiol*, 21, 78–89.

Appendix

Appendix A. Response sheet for subjective ratings.

Note that the first three questions were answered twice, once for program A and once for program B.

While listening to the conversation in noise, the background noise is

5 Not noticeable
4 Somewhat noticeable
3 Noticeable but not intrusive or distracting
2 Fairly conspicuous, somewhat intrusive or distracting
1 Very conspicuous, very intrusive or distracting

While listening to the conversation in noise, the target talker is

5 Very clear
4 Rather clear
3 Somewhat clear
2 Rather unclear
1 Very unclear

While listening to the conversation in noise, the general quality feels

5 Very natural
4 Fairly natural
3 Somewhat natural
2 Fairly unnatural
1 Very unnatural

My overall rating of device A compared to device B is

3 A is much better
2 A is better
1 A is slightly better
0 They are about the same
−1 B is slightly better
−2 B is better
−3 B is much better