Loss of speech perception in noise – causes and compensation

JORGE MEJIA[1,2], HARVEY DILLON[1,2,*], RICHARD VAN HOESEL[1,4],
ELIZABETH BEACH[1,2], HELEN GLYDE[1,2], INGRID YEEND[1,2], TIM BEECHEY[1,2],
MARGO MCLELLAND[1,2], ANNA O'BRIEN[1,4], JÖRG BUCHHOLZ[1,2,3],
MRIDULA SHARMA[1,3], JOAQUIN VALDERRAMA[1,2], AND WARWICK WILLIAMS[1,2]

[1] *The HEARing Cooperative Research Centre, Australia*

[2] *National Acoustic Laboratories, Macquarie University, Sydney, Australia*

[3] *Department of Audiology, Macquarie University, Sydney, Australia*

[4] *Department of Audiology and Speech Pathology, The University of Melbourne, Melbourne, Australia*

Any damage within the cochlea, whether affecting hearing thresholds or high threshold nerve fibres, that affects the resolving power of the cochlear, necessitates a higher input signal-to-noise ratio to achieve normal speech understanding in noise. Other than wireless remote microphone systems, super-directional beamformers are the most effective way to achieve this. To optimise their performance, they should have beam widths that are neither too narrow nor too broad, attenuate off-beam signals in a way that preserves spatial awareness of the environment, and adapt to changing competing signals fast enough to suppress them but not so fast as to distort the target signal. This paper reports on the advantages and limitations of super-directional beamformers as measured in six different experiments.

INTRODUCTION

It is now well established in animal studies that high levels of noise, even for a few hours, can damage the auditory system in ways that are not evident in the audiogram. In particular, high threshold, low spontaneous rate, afferent fibres originating at inner hair cells are destroyed, starting with destruction of the synapse within days of the noise exposure (Furman *et al.*, 2013; Kujawa and Liberman 2009). There is some uncertainty about how this finding translates to humans, and if so, what the consequences for humans are. The first part of this paper shows the context in which this question is being comprehensively investigated. A likely consequence is that some people with little or no elevation in hearing thresholds require a better signal-to-noise ratio (SNR) to communicate than do their peers with the same hearing thresholds (Plack *et al.*, 2014).

The second part of the paper very briefly summarises one reason, which turns out to be simple audibility, why people with elevated hearing thresholds also require a better SNR than people with normal hearing. Although the reason may be simple, the solution is not, as there is a limit to how much amplification a person with hearing loss will tolerate. We are not yet at the stage of being able to analyse precisely why, other than inadequate audibility, damage to the hearing system creates difficulties in recognising and

---

*Corresponding author: harvey.dillon@nal.gov.au

Proceedings of ISAAR 2015: Individual Hearing Loss – Characterization, Modelling, Compensation Strategies. 5th symposium on Auditory and Audiological Research. August 2015, Nyborg, Denmark. Edited by S. Santurette, T. Dau, J. C. Dalsgaard, L. Tranebjærg, and T. Andersen. ISBN: 978-87-990013-5-4. The Danavox Jubilee Foundation, 2015.

understanding speech for an individual, especially in noisy situations. We are even further from being able to build inverse processes (in the unlikely event that is even possible) into hearing devices to restore normal functioning. It seems extremely likely, however, that anyone with a hearing problem, whatever its underlying origin, will benefit from devices that provide them with a better SNR than they would have access to without any device. The final, and major, section of the paper therefore provides an overview of a series of experiments designed to evaluate a novel method of improving SNR. The method is based on a binaural beamformer that provides a greater degree of directivity than hearing aids working in isolation on each side of the head can provide.

NOISE EXPOSURE, SPEECH RECOGNITION, COGNITION, AND COCHLEAR FUNCTIONING

Our primary interest is in understanding the relationship between noise exposure, cochlear functioning, and the consequences of the latter for speech recognition. Our hypothesis is that any relationship between noise exposure and speech recognition will be completely mediated by the effect that the noise exposure has had on cochlear functioning, and of course its downstream effects on auditory nerve fibres and higher centres (Bramhall *et al.*, 2015; Schaette and McAlpine, 2011). Speech recognition, however, is very likely to be affected by cognitive abilities (Helfer and Jesse, 2015). It is also possible that it is affected by musical training or experience, either by improving auditory brainstem functioning (Skoe and Kraus, 2013; Slater *et al.*, 2015) or by improving cognitive abilities, so consequently we need to measure these as well.

Figure 1 shows the relationships that we are hypothesising may exist between the quantities measured. Lifetime noise exposure, estimated from a questionnaire (Beach *et al.*, 2013) is assumed to damage high-threshold nerve fibres, outer hair cell (OHC) functioning, and possibly low-threshold nerve fibres. The latter two forms of damage (along with any reduction in stria vascularis effectiveness that affects their functioning) are presumed to determine hearing thresholds. OHC damage should be observable in the levels of otoacoustic emissions both transient (TEOAE) and distortion product (DPOAE). High threshold fibre damage should be observable behaviourally in the detection of tones in threshold equalizing noise (TEN test; Moore *et al.*, 2012), in reduced sensitivity to temporal fine structure (TFS; Moore and Sek, 2009), and in elevated thresholds for detection of amplitude modulation (AM). In the latter two tests, lower level background noise is used to limit the ability of low and medium threshold fibres to contribute to the task. Damage to high threshold fibres should also be observable as a reduced growth of envelope following response as modulation depth increases (Bharadwaj *et al.*, 2014), reduced amplitude of wave I in a click ABR (Schaette and McAlpine, 2014; Stamper and Johnson, 2014), and a decreased magnitude and coherence of a speech ABR (Anderson *et al.*, 2013). Stimuli for these electrophysiological tests will also be masked to maximise sensitivity to high threshold fibre activity.

The three cochlear variables are hypothesised to affect speech recognition, whether measured behaviourally with the Listening in Spatialised Noise Sentences test (LiSN-S; Cameron and Dillon, 2007) or via self report with the Speech Spatial Qualities test (SSQ12; Noble *et al.*, 2013). Each of the measures of speech recognition may be affected by verbal memory, attention (the Test of Every Day Attention;

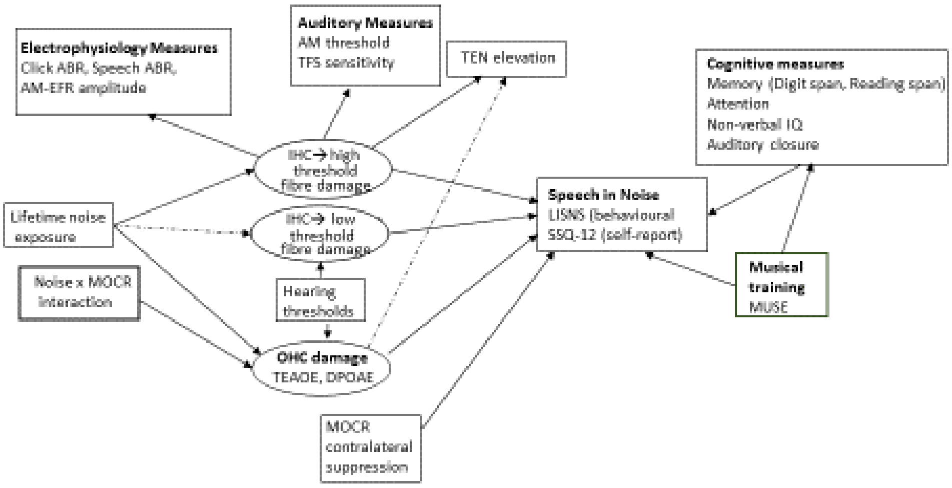

Fig. 1: Hypothesised relationships between factors affecting the perception of speech in noise. The diagram shows latent variables (that cannot directly be observed) as ellipses, and indicators of those variables or other measurable quantities as rectangles.

Robertson *et al.*, 1996), non-verbal intelligence and auditory closure ability. Verbal working memory can be assessed with digit span forward and reverse, and the Reading Span Test (Daneman and Carpenter, 1980). Musical experience can be estimated with the Music Use (MUSE) questionnaire (Chin and Rickard, 2012), which provides indices relating to both level of musical training and experience in playing music. Finally, the medial olivo-cochlear response (MOCR) is assumed to assist in recognising speech in noise, and to help protect OHCs against noise damage, hence the box showing an interaction between MOCR strength and noise exposure.

At the time of writing, behavioural data has been measured on 78 adults aged 30 to 55 years with hearing thresholds in the normal or 'near to normal' range. Their noise exposure varied greatly, up to an estimated 62,000 Pa^{2}hrs. Musical training and experience likewise varied over a wide range. Electrophysiological data have so far been obtained on only 12 participants. Analysis of the results will be reported on in later publications, when further behavioural and electrophysiological data have been collected.

HEARING IMPAIRMENT AND SPEECH PERCEPTION IN SPATIALLY SEPARATED COMPETITION

Investigation into the impact of hearing impairment on speech recognition in spatially separated distractors further demonstrates the need for improved SNRs. Glyde et al (2013a) tested 80 people, aged 7-89 years with hearing levels ranging from normal to moderately-severe, on the LiSN-S and found increasing hearing impairment correlated with worsening speech reception thresholds in noise (SRTn) (see Fig. 2). This relationship existed despite the use of NAL-RP amplification. The deficit was strongest in the test conditions in which the target speech was spatially separated from the distractors due to decreasing spatial release from masking (SRM) with increasing hearing loss (partial $r^2 = 0.66$).

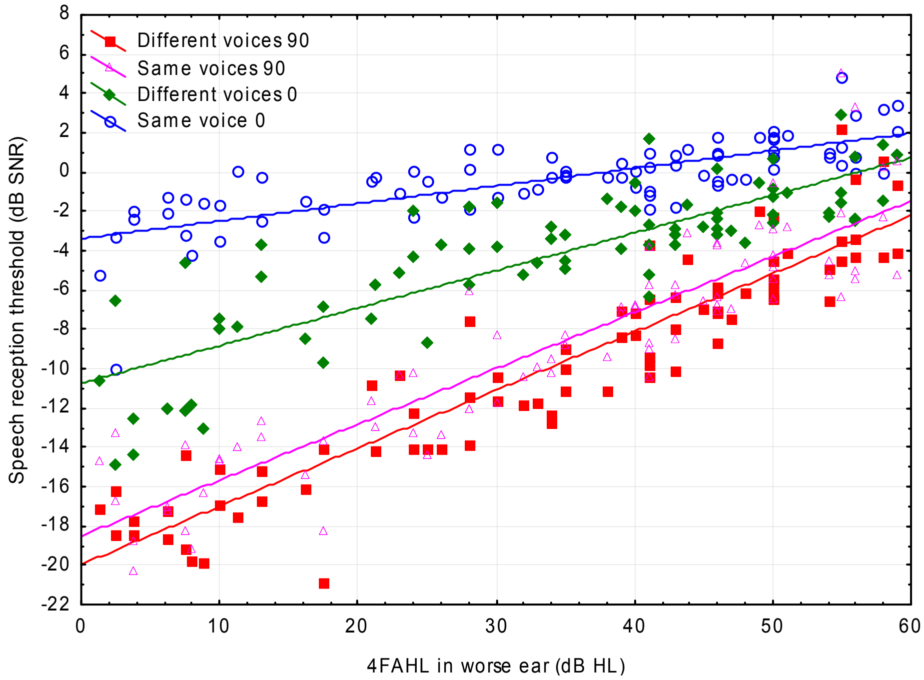

Fig. 2: Variation of SRTn measured with the LiSN-S test as four-frequency average hearing thresholds in the worse ear vary from normal to 60 dB HL. The four conditions of the LiSN-S comprise the talker being the same or different voice as the distractors, combined with the distractors being at the same (0°) location as the talker or at different (±90°) locations. Reprinted with permission from Glyde et al (2013a).

The underlying cause of this apparent loss of ability to use spatial cues was investigated in a subsequent series of experiments. Firstly, by creating versions of the LiSN-S test stimuli which contained only interaural level differences (ILDs) or interaural time differences (ITDs), and comparing normal-hearing adults' performance on these versions to performance with both cues available, it was ascertained that ILDs alone provided as great SRM as the two cues together (Glyde *et al.*, 2013b). This result suggested that ILD interpretation or transmission was the most likely barrier to achieving SRM. Given ILD's dominance in the high frequencies, limited audibility of the SNR benefits arising from ILDs could explain the results shown in Fig. 2.

This hypothesis was examined in Glyde *et al.* (submitted) where frequency-specific filtering was applied to the stimuli so that sensation levels were matched between a sample of normal-hearing and hearing-impaired adults. Speech reception thresholds were compared at three amplification levels (NAL-RP, NAL-RP+25%, NAL-RP+50%). Increased amplification significantly improved SRM ($p< 0.001$). Therefore if better audibility could be provided to hearing-impaired individuals, better performance in spatially separated competition is expected. However, high-frequency gain considerably in excess of that provided by NAL-RP would be needed to enable close to normal SRM, and this much high-frequency gain is generally not acceptable to hearing aid wearers, and is often not possible because of feedback oscillation.

IMPROVING SNR THROUGH BINAURAL BEAMFORMING

The most promising method for improving SNR is to use directivity to provide greater amplification for sounds coming from a target direction than to sounds coming from other directions. Directional microphones mounted within a hearing aid are very limited in the extent to which they can do this, as the sounds arriving at two closely spaced ports differ very little in either time of arrival or level. Signals arriving at one side of the head, however, differ greatly in both time of arrival and level from signals arriving at the other side, for sources other than those directly in front or behind the listener. Inputs to the beamformer applied to each ear come from the output of conventional directional microphones. Essentially, at any moment in time, the beamformer gives maximum amplification to those frequency components of the signal that have the same level and phase (i.e., time of arrival) at the two sides of the head, and progressively less amplification to those components that differ in either amplitude or phase. Many variations are possible while still conforming to this general principle. Variations include:

- The azimuth variation from straight ahead beyond which sounds are attenuated (and hence the target beam width);
- The degree to which off-beam signals are attenuated;
- The rate at which the characteristics of the beamformer are allowed to adapt, and the frequency resolution with which the characteristics are determined;
- The extent to which the original time and level differences at each ear are retained in the outputs fed to each ear (and hence the extent to which spatial awareness is retained);
- The relative reliance placed on inter-aural time differences versus inter-aural level differences in determining the weights given to each component; and
- The way in which each of the above considerations is varied with frequency.

The results reported in this paper were obtained with beamformers that were progressively improved by fine tuning these variations over several years to optimise the combination of SNR enhancement, lack of perceptible distortion, and retention of spatial information. The choices that affect each of these also affect the other two desired characteristics, so optimising the trade-off is necessary. An audio-visual example of the performance that is possible with such beamformers can be accessed at www.hearingcrc.org/xc/xc4-applications-of-binaural-signal-processing/.

Describing beamformer performance

Unlike a conventional, static directional microphone, the CRC beamformers are adaptive, so do not have a single polar pattern or directivity index that captures their performance. Figure 3 shows how dramatically the polar diagram can change when other signals are present in addition to the target signal, the sensitivity to which the polar diagram represents.

Beam width

Just how super-directional should a beamformer be? The narrower the beam-width, the greater the SNR enhancement that is possible, especially when the dominant competing sound(s) come from the frontal hemi-field. However, the narrower the beam-width, the greater the chance that listeners will misalign their heads, thus decreasing sensitivity to the target, or that targets will be distorted if, due to the effects

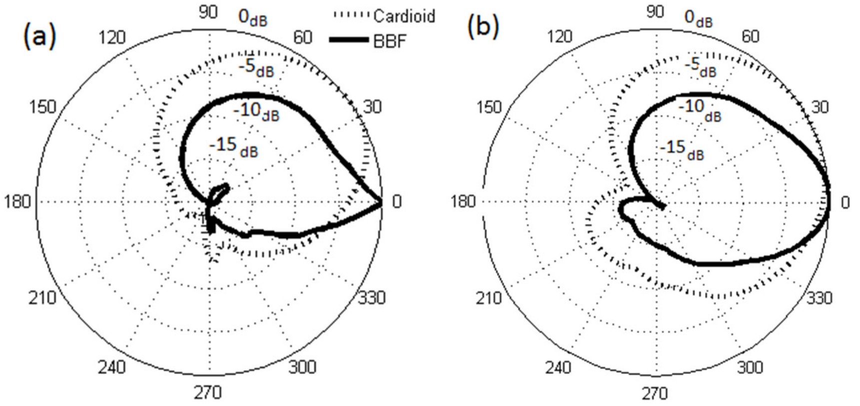

Fig. 3: Polar diagrams measured for a beamformer (a) when a single target source is varied in azimuth, and (b) when in addition to the target signal, speech babble comes from eight loudspeakers spaced every 5 degrees around the listener in the horizontal plane. In both diagrams, the dotted line shows the pattern for a conventional directional microphone as a comparison.

of other signals, some target components are assessed as coming from within the beam aperture and others as coming from outside the beam aperture.

Figure 4 shows, for three different beam-widths, how the beamformer improved intelligibility relative to two independent cardioid directional microphones for seven listeners with mild to moderate hearing loss. Each listener was tested at the SNR for which he or she obtained 50% of items correct when listening to the cardioid microphones. The narrowest beamformer gave the worst performance for those listeners most able to communicate at very poor SNRs. The remaining results in this paper were therefore obtained using beamformers that did *not* have extremely narrow beams.

Retention of spatial cues

Spatial cues are important to listeners for many reasons: awareness of one's surroundings, localization of desired target sounds, and spatial separation of desired targets from unwanted competition. To investigate their effect on intelligibility, spatial cues were intentionally removed from both the beamformer and the reference condition (independent cardioid directional microphones). In both cases, the signals normally applied separately to each ear were mixed and the mixture applied to both ears – that is, diotic presentation.

Figure 5 shows the results. When spatial cues were present for both processing types, beamformer performance was 17 percentage points higher than cardioid performance (significant with p=0.003). Removal of spatial cues from the cardioid microphone outputs decreased performance by 39 percentage points. By contrast removal of the spatial cues from the beamformer decreased its performance by only 8 percentage points. The interpretation of this in unclear. One possibility is that the beamformer processing was already taking advantage of spatial cues, much as listeners do when listening to separate left and right ear signals, so that a loss of spatial cues is of less consequence for intelligibility. A second possibility is that the beamformer was not adequately retaining spatial cues in the

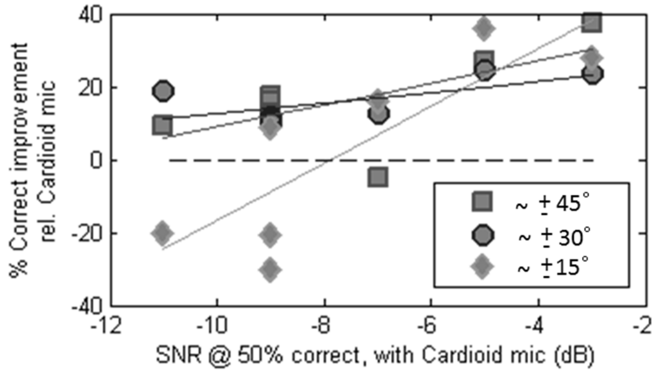

Fig. 4: Intelligibility improvement (percentage points) relative to independent cardioid microphones for beamformers with different beam widths. The target was presented to the front, and independent speech babble was presented from each of the remaining 45-degree intervals around the listener.

dichotic condition, so there were fewer spatial cues to remove. Our interpretation, based on careful but informal listening trials was that both of these possibilities were occurring, and we strengthened the retention of spatial cues within the beamformer for subsequent experiments.

Dynamic listening situations

Dynamic listening situations, where the direction of arrival of the target sound changes rapidly, such as in a group discussion, are potentially challenging for beamformers. To investigate this, we compared performance for a single frontal talker to performance with two talkers engaged in natural conversation. In the latter case, one talker was presented from the front and the second was randomly presented from either −45° or +45°. Listeners were encouraged to turn towards each talker throughout the conversation. Strong competing talkers were included at −45° and +135°, or in a second configuration at −90° and +90°. In both cases, weaker background (uncorrelated) cafeteria noises were placed at all other multiples of 45° around the circle. In this experiment, rather than measure speech intelligibility, we measured the acceptable noise level (ANL) for cardioid microphone and for beamformer processing. Listeners first adjusted the gain for the target sound to give a comfortable level, and then the competing sounds to the loudest level they would be willing to tolerate for sustained listening. The ANL was the SNR at this just-tolerable noise level.

Figure 6 shows the improvement in ANL offered by the beamformer over the independent cardioid microphones for 4 listeners with normal hearing and 11 listeners with mild to moderate sensorineural hearing loss. For reasons that will become apparent later in this paper, we think that the beamformer offers the greatest advantages to those with the greatest hearing loss. For the two-talker condition, however, there appears to be less advantage for the listeners with a mild loss than for those with either normal hearing or moderate loss. We therefore fitted a quadratic curve to the data. For the single frontal talker condition, the advantage of the beamformer varies less markedly, but again a quadratic curve was fitted. The advantage is, nonetheless, about 2 dB, almost independent of hearing loss over this range of hearing losses. This 2-dB improvement in ANL enabled by the beamformer is smaller than we have obtained in other single talker experiments, a difference we ascribe to the strong competition being only 45° away from the target talker in this experiment, rather than being equally distributed across azimuths. A possible interpretation of the quadratic variation of benefit in the two talker conversation is that:

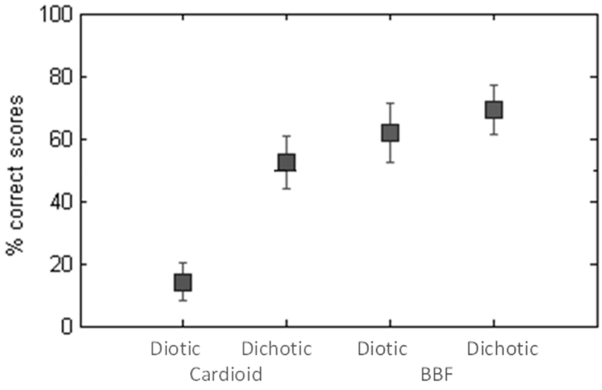

Fig. 5: Percent correct intelligibility, with 95% confidence intervals for independent cardioid directional microphones and the binaural beamformer under dichotic and diotic conditions. Target was presented to the front, and independent two-talker noise was presented from each of the remaining 45 degree intervals around the listener.

- those with normal hearing had sufficiently good hearing to quickly and accurately track the talker location and hence orient their head optimally, even if the salience of localization cues was reduced by the beamformer;
- those with mild loss had their ability to track the target talker negatively impacted by the beamformer; which the improved SNR offered by the beamformer only just made up for; and
- those with moderate loss had reduced ability to track the target talker even with the cardioid microphones, and so were less affected by the reduced spatial cues in the beamformer.

Real-life noises, reverberation and distances

The performance of all directional microphones is adversely affected by increasing reverberation times and distance from the source, as directivity cannot be useful if effectively all sounds come from all directions. Evaluating beamformers in real-world conditions is therefore important to get a proper view of their capabilities. To achieve this in a controlled manner, recordings of the background sounds picked up by dual omnidirectional microphones inside behind-the-ear (BTE) hearing aid cases worn on each side of the first author's head were made in 30 real-world locations. At each location, the impulse response from an imaginary talker's position to each of the four microphones was also recorded. These impulse responses were later convolved with anechoic speech to provide the target talker signal that would have been received in each of these situations had a talker been present at the appropriate distance directly in front of the listener. Target to background signal to noise ratios were set appropriate to the actual SPL of the background noise based on the data in Pearsons *et al.* (1977). The combined target and background noise signals were then processed to provide stereo signals corresponding to:

- omnidirectional microphones;
- cardioid directional microphones;
- binaural beamformer, with retention of some spatial information;
- an "ideal" beamformer, formed by using the cardioid directional microphones and simply then increasing the SNR by 5 dB.

Listeners (12 with normal hearing and 24 with hearing loss) were asked to rate, using a slider scaled from 0 (very poor) to 1 (perfect), each listening situation for listening effort, naturalness, noisiness, smoothness, distortion and overall acceptability.

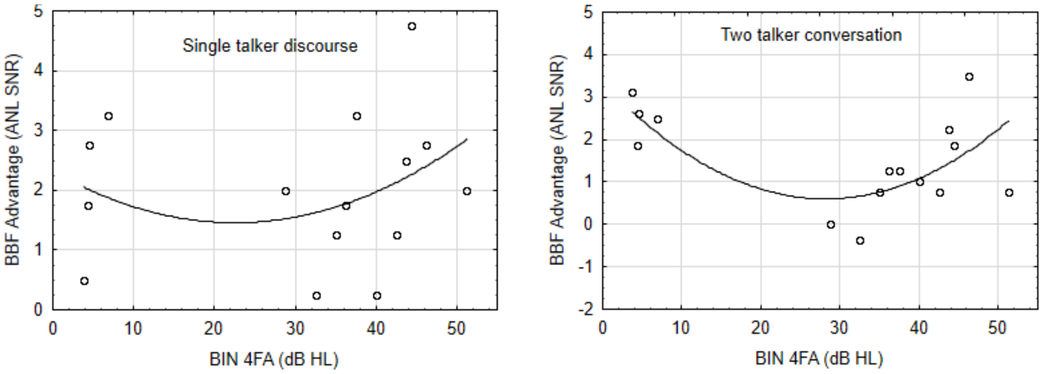

Fig. 6: Improvement in ANL for the beamformer relative to cardioid microphones in the one-talker and two-talker situations.

Figure 7 shows the sound quality rating differences relative to the cardioid. On all measures beamformer processing was preferred to cardioid processing. The extent of the preferences varied from 0.15 scale points (distortion) to 0.55 scale points (listening effort). In each case, preference for the beamformer was similar to that for the ideal beamformer, indicating that the beamformer *sounded* like it was giving about a 5-dB improvement in SNR. One of the benefits that was evident for the beamformer was a marked reduction in wind noise in those outdoor situations where wind noise was present. Although the omni microphone was rated below the cardioid microphone for 5 out of the 6 qualities, the difference is always small. This reflects the very limited advantage that a standard directional microphone can provide in reverberant listening situations. The higher directivity obtainable with a beamformer substantially increases the range of situations in which directivity is beneficial.

Application to cochlear implants

It seems likely that the net benefit offered by beamformers reflects the advantage achieved by increasing the SNR, offset by the disadvantage caused by any loss of spatial information and any distortions created by the beamformer that are perceived by the listeners. Because of the limited auditory ability of listeners with severe loss, including those listening through cochlear implants, these disadvantages should be smaller, thus creating a larger net benefit for these listeners. Performance of the beamformer was evaluated for 10 users of bilateral cochlear implants, under conditions of sparse competition (competing talkers at 60°, 90°, and 270°) and diffuse competition (competing talkers at 45°-intervals from 45° to 315°).

Figure 8 shows the SRTn values achieved with the beamformer. Depending on performance with the omni mic, the improvement in SRTn was on average 8.8 dB SNR for the sparse talker condition, and varied from 4 to 8 dB for the diffuse competing talker condition. Although it was not possible to use a cardioid reference condition in this experiment, the benefit relative to omni microphones considerably exceeds the benefit in SRTn typically offered by cardioid microphones relative to omni.

SRTn benefit at positive SNRs

The results so far, especially in combination with the subjective impression of beamforming, contain a paradox. The subjective impression is of a very marked

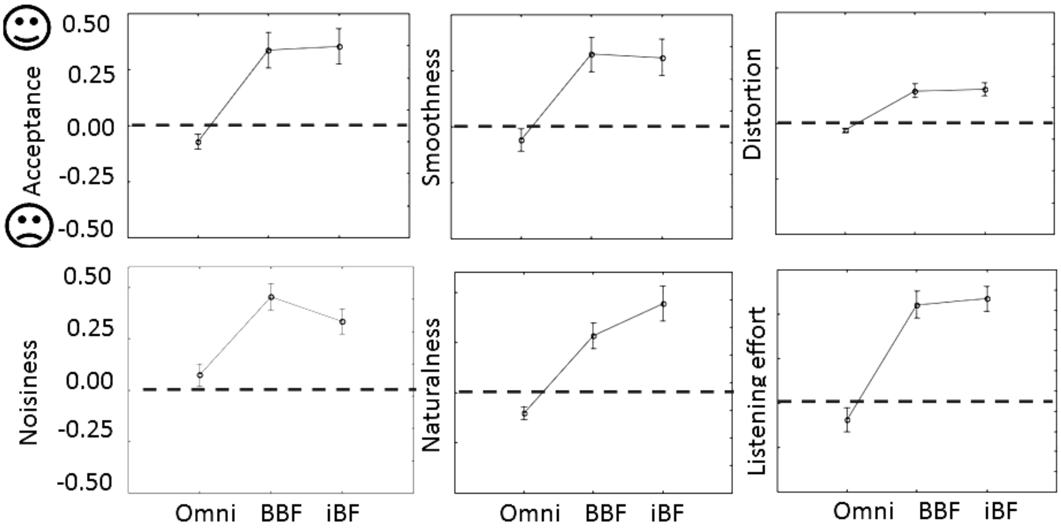

Fig. 7: Sound quality ratings relative to cardioid for the omni, beamformer (BBF) and ideal beamformer (IBF) microphone systems, averaged across the 30 listening situations.

improvement relative to cardioid microphones, which is consistent with the SRTn improvement for cochlear implantees (Fig. 8) and the quality ratings for hearing aid wearers (Fig. 7). The intelligibility improvements re cardioid at SRTn for hearing aid wearers are, however, "only" around 20 percentage points (Figs. 4 and 5), equivalent to about a 2-dB improvement in SRTn to cardioid microphones at 0 dB SNR, decreasing to 3 dB at −15 dB SNR. Is the smaller benefit measured in SRTn because SRTn typically occurs at very negative SNRs, or is it because the improved SNRn is offset by some distortions introduced by the beamformer, such as a reduction in the salience of spatial cues?

To investigate this, we performed an additional experiment in which we made the test material difficult by using casually articulated nonsense CVC syllables, and in which we targeted the 50% point on the psychometric function which was 20% lower than the scores obtained in quiet. This was evaluated in a diffuse background formed from competing talkers at 45° intervals from 45° to 315°.

Figure 9 shows the SRTn and acceptance scores for beamformer relative to cardioid. As shown in the figure, although we created the speech test with the aim of hearing impaired subjects obtaining SRTn at SNRs at or above 0 dB (typical of real-life conversational levels), SRTn with the cardioid microphone nonetheless ranged from -10 to +2 dB across the 26 participants with mild to moderate hearing loss. Beamformer SRTn benefit relative to cardioid was, on average, 1.8 dB SNR. Participants also rated acceptability of the amplified sound on a 1 to 10 scale, when measured at a SNR of 0 dB. On average, the beamformer produced a score 1.5 scale points higher than the cardioid microphone.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the Australian Department of Health, the Cooperative Research Centre scheme, the National Health and Medical Research Council, and the Hearing Industry Research Consortium.

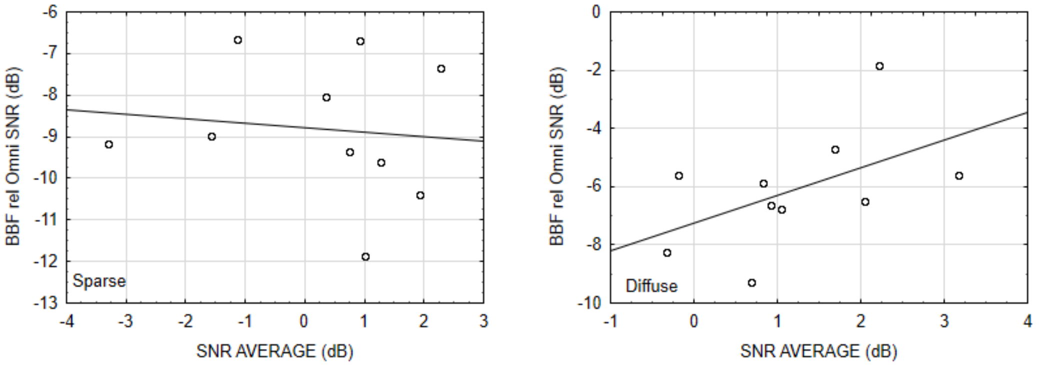

Fig. 8: SRTn in noise (individual data points and regression lines) for the beamformer SRTn minus omnidirectional microphone SRTn versus performance averaged across the SRTn values for BBF and omnidirectional microphone.

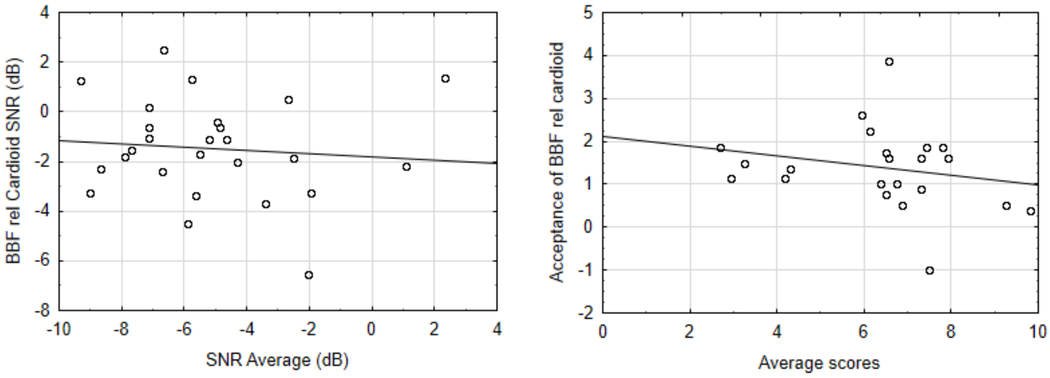

Fig. 9: SRTn performance of BBF relative to cardioid for each listener relative to the scores averaged across BBF and cardioid. (a) shows difference in SRTn and (b) shows difference in acceptance ratings. The solid lines show the corresponding regressions.

REFERENCES

Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (**2013**). "Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance," J. Speech Lang. Hear. Res., **56**, 31-43.

Beach E., Gilliver, M., and Williams, W. (**2013**). "Sound Check Australia: A citizen science approach to noise and hearing conservation research," National Hearing Conservation Association Conference, St Petersburg, Florida.

Bharadwaj, H.M., Verhulst, S., Shaheen, L., Liberman, M.C., and Shinn-Cunningham, B.G. (**2014**). "Cochlear neuropathy and the coding of supra-threshold sound," Front. Sys. Neurosci., **8**, 26.

Bramhall, N., Ong, B., Ko, J., and Parker, M. (**2015**). "Speech perception ability in noise is correlated with auditory brainstem response wave I amplitude." J. Am. Acad. Audiol., **26**, 509-517.

Cameron, S. and Dillon, H. (**2007**). "Development of the Listening in Spatialized Noise – Sentences Test (LISN-S)." Ear. Hearing, **28**, 196-211.

Chin, T. and Rickard, N. (**2012**). "The music USE (MUSE) questionnaire: An instrument to measure engagement in music." Music Percept., **29**, 429-446.

Daneman, M. and Carpenter, P. (**1980**). "Individual differences in working memory and reading." J. Verb. Learn. Verb. Be., **19**, 450-466.

Furman, A.C., Kujawa, S.G., and Liberman, M.C. (**2013**). "Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates," J. Neurophysiol., **110**, 577-586.

Glyde, H., Cameron, S., Dillon, H., Hickson, L., and Seeto, M. (**2013a**). "The effect of hearing impairment and aging on spatial processing ability." Ear. Hearing, **34**, 15-28.

Glyde, H., Buchholz, J.M., Dillon, H., Cameron, S., and Hickson, L. (**2013b**). "The importance of interaural time differences and level differences on spatial release from masking." J. Acoust. Soc. Am. **134**, EL147-152.

Glyde, H., Buchholz, J.M., Nielsen, L., Best, V., Dillon, H., Cameron, S., and Hickson, L. (**submitted**). "Effect of audibility on spatial release from speech-on-speech masking."

Helfer, K.S. and Jesse, A. (**2015**). "Lexical influences on competing speech perception in younger, middle-aged, and older adults." J. Acoust. Soc. Am., **138**, 363-376.

Kujawa, S.G. and Liberman, M.C. (**2009**). "Adding insult to injury: Cochlear nerve degeneration after "temporary" noise-induced hearing loss." J. Neurosci., **29**, 14077-14085.

Moore, B.C.J. and Sęk, A. (**2009**). "Sensitivity of the human auditory system to temporal fine structure at high frequencies." J. Acoust. Soc. Am., **125**, 3186-3193.

Moore, B.C.J., Creeke, S., Glasberg, B.R., Stone, M.A., and Sęk, A. (**2012**). "A version of the TEN Test for use with ER-3A insert earphones." Ear. Hearing, **33**, 554-557.

Noble, W., Jensen, N.S., Naylor, G., Bhullar, N., and Akeroyd, M.A. (**2013**). "A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12," Int. J. Audiol., **52**, 409-412.

Pearsons, K.S., Bennett, R.L., and Fidell, S. (**1977**). "Speech levels in various noise environments." Washington, D.C., U.S. Environmental Protection Agency.

Plack, C.J., Barker, D., and Prendergast, G. (**2014**). "Perceptual consequences of "hidden" hearing loss," Trends Hear., **18**, 1-11.

Robertson, I.H., Ward, T., Ridgeway, V., and Nimmo-Smith, I. (**1996**). "The structure of normal human attention: The Test of Everyday Attention." J. Int. Neuropsychol. Soc., **2**, 525-534.

Schaette, R. and McAlpine, D. (**2011**). "Tinnitus with a normal audiogram: Physiological evidence for hidden hearing loss and computational model." J. Neurosci., **31**, 13452-13457.

Skoe, E. and Kraus, N. (**2013**). "Musical training heightens auditory brainstem function during sensitive periods in development." Front. Psychol., **4**, 622.

Slater, J., Skoe, E., Strait, D.L., O'Connel, S., Thompson, E., and Kraus, N. (**2015**). "Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program." Behav. Brain Res., **291**, 244-252.

Stamper, G.C. and Johnson, T.A. (**2014**). "Auditory function in normal-hearing, noise-exposed human ears," Ear Hearing, **36**, 172-184.