Unilateral crosstalk cancellation in normal hearing participants using bilateral bone transducers

Robert W. J. Mcleod and John F. Culling,

School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, CF10 3AT, U.K.

Date: Friday, 19 June 2020

Running title: Bone-conducted sound.

PACS numbers: 43.66.Ts, 43.66.Pn

Correspondence address: -

Robert Mcleod

School of Psychology,

Cardiff University,

Tower Building, Park Place,

Cardiff,

CF10 3AT

U.K.

mcleodrw@cf.ac.uk

Abstract

It is possible to psychophysically measure the phase and level of bone conducted sound at the cochleae using two bone transducers (BTs) [Mcleod & Culling, *J. Acoust Soc. Am.* 146, 3295–3301 (2019)]. The present work uses such measurements to improve masked thresholds by using the phase and level values to create a unilateral crosstalk cancellation system. To avoid changes in the coupling of the BT to the head, testing of tone and speech reception thresholds with and without crosstalk cancellation had to be performed immediately following the measurements without adjustment of the BT. To achieve this, a faster measurement method was created. Previously measured phase and level results were interpolated to predict likely results for new test frequencies. Testing time to collect the necessary phase and level values was reduced to approximately 15 min by exploiting listeners' previous measurements. The inter-cochlear phase difference and inter-cochlear level difference were consistent between experimental sittings in the same participant but different between participants. Addition of a crosstalk cancellation signal improved tone and speech reception thresholds for tones/speech presented with one BT and noise presented on the other by an average of 12.1 dB for tones and 13.67 dB for speech.

I. Introduction

Few studies have investigated the benefits of bilateral bone-conduction hearing aids. Using sound field measurements, improvements of 2-15 dB in masked tone thresholds compared to unilateral fitting have been demonstrated for adult listeners (Bosman et al.; Priwin et al., 2004). Speech reception thresholds in quiet have improved by 4.2 dB (Bosman et al., 2001). However, these benefits may be purely due to amplification from two hearing aids rather than increased ability to process sound binaurally. In order to investigate true binaural processing advantages Binaural Masking Level Differences (BMLDs) have been used. These have shown significant benefit (6-6.1 dB) at low frequencies (125-500 Hz), but no significant benefit at 1000 Hz (Bosman et al., 2001; Priwin et al., 2004). Sound localisation and lateralization judgements have also been shown to improve significantly (Bosman et al., 2001). This shows that there is a true binaural advantage although it is severely limited compared to normal hearing due to crosstalk within the head (Deas et al., 2010).

Crosstalk cancellation was originally conceived by Bauer (1961) in order to more accurately reproduce binaural recorded signals from two loudspeakers. The technique was later put into practice by Schroeder and Atal (1963). Several different methods of crosstalk cancellation have been developed. However, they all attempt to implement the theoretical “ideal crosstalk cancellation” taking into account real world limitations such as the dynamic range of the amplifier or transducer. This is problematic because ideal crosstalk cancellation has the potential to require high output levels in order to cancel crosstalk when the two direct signals are close to being in phase at the receivers. This problem arises because destructive interference will occur to a large proportion of the desired signal. In this ‘ill-condition,’ where the signal phases are close, it can leave the system very prone to small measurement inaccuracies as well as head movement. Thus, at frequencies where there is little interaural phase difference, crosstalk cancellation cannot be achieved reliably. For bone transducers located on either side of a human head, these small phase differences occur mostly at low frequencies.

For frequencies above about 1 kHz, Mcleod and Culling (2019) demonstrated the equivalence of two measurement techniques; the phase and level measured by cancelling the signal from one bone transducer (BT) using another gave equivalent phase and level results when compared to cancelling each separately using sound presented over earphones. In the present study, we introduce a faster method of measuring the phase and level results necessary for crosstalk cancellation and show that the resulting crosstalk cancellation can be used to substantially reduce masking through improved stereo separation.

II. Experiment 1

The first experiment took initial measurements of the phases and amplitudes required for crosstalk cancellation at each ear. These baseline measurements for each participant were used in Exps. 2 and 3 to facilitate rapid remeasurement prior to testing the effectiveness of crosstalk cancellation in masked threshold tasks. The methodology was approved by Cardiff University Psychology Department Ethics Committee.

A. Methods

1. Equipment

Sound presentation and data calculation was performed with the use of MATLAB®. A USB ESI MAYA44 USB+ four-channel digital-to-analog converter was used in conjunction with an 8-channel Behringer Powerplay Pro-8 Headphone amplifier to pass audio signals to two B71 (Radioear) BTs. A pair of Etymotic ER2 insert earphones with ER1-14B eartips were inserted into the ears of the participants to prevent air-borne sound radiated from the BTs from interfering with the crosstalk cancellation results. ER2s were used rather than ear plugs for consistency with previous work but were not used to present sound. BT placement was the same as outlined in Mcleod and Culling (2017, 2019); BTs were attached to a pair of spectacle frames and pressed against the head using a softband. There was no adjustment of the BT positioning once measurements of phase and level had begun. All testing was performed in a single-walled Industrial Acoustics Company (IAC) sound-attenuating booth within a sound-treated room. A computer screen was visible outside the booth window with a keyboard and mouse inside the booth for participants to adjust phase and level differences as well as input transcripts in Exp. 3.

2. Stimuli

The stimuli were pairs of sinusoids of the same frequency, but adjustable phase and level. presented via different bone transducers.

3. Participants

Three participants aged between 21 and 29 years old were recruited from Cardiff University and were paid for each testing session. All had previous experience with psychoacoustic experiments, were native English speakers and had self-reported normal hearing with no previous history of ear pathology. Otoscopic examination prior to testing was normal and ensured that wax levels were low enough to safely use deeply inserted tubephones. Pure-tone audiometry was considered unnecessary, because there was no expectation that any mild cochlear hearing loss would interact with the required measurements. All participants had performed at least 5 hours of testing using tone-cancellation tasks in other experiments prior to data collection.

4. Procedure

The procedure for measuring phases and levels required for crosstalk cancellation were previously described as the 'two-BT' method by Mcleod and Culling (2019). The two-BT method was used here because it is readily applicable to the target population of patients with severe bilateral conductive loss. In this technique, the phase and level of a tone at one BT is adjusted in order to cancel the signal from the contralateral BT at the ipsilateral cochlea. Perceptually, the task is to maximize the laterality of the tone by adjusting two controls. A limitation of this method is that it cannot be performed at frequencies below about 1 kHz due to the interaction of interaural time and level cues (see General Discussion), but, as noted in the Introduction, crosstalk cancellation is difficult to achieve at low frequencies in any case.

Participants underwent five trials on each side and at each frequency in order to obtain a set of initial phase and level data. A prediction algorithm was used to aid the method of adjustment. It placed the stimuli as close as possible to a predicted match at the beginning of a trial. Adjustments were thus only made to refine these predictions.

In the adjustment task, participants cancelled a pure tone at one frequency at the target ear by adjusting the phase and level of a contralaterally presented tone, resulting in a strongly lateralized percept. Once achieved, participants could then change the frequency by multiples of 20 Hz using a mouse scroller. When the frequency is changed the laterality is reduced somewhat because the required phase and level are a little different. The participant would make further adjustments to the phase and level difference in order to increase the laterality and thus cancel the tone at one ear for the new frequency. Keeping the phase and level values from one frequency to the next is advantageous, because the phase and level needed for cancellation only needs to be varied by a small amount to optimize the cancellation rather than starting from an unknown point.

The starting frequency was 3 kHz. If participants could not cancel sound at this frequency, then the frequency was increased by 200 Hz until cancellation was possible.

Participants were unable to achieve cancellation at the start frequency of 3 kHz on two occasions, but after successful cancellation at other frequencies were able to reattempt and cancel 3 kHz. Cancellation was possible on further testing because phase and level results for frequencies close to the target frequency better informed the starting point for the search. Once an initial crosstalk cancellation result had been achieved, the participant increased the presentation frequency by 200 Hz and again attempted crosstalk cancellation. During this process, the values of level and phase difference as well as the frequency were displayed on the screen. Participants were told that in most cases an increase in frequency would result in an increase in phase difference. A further iteration of increasing the frequency by 200 Hz and keeping the previous phase and level difference settings was performed. Once the cancellation program had at least three phase and level results from different frequencies it could start predicting the phase and level needed for cancellation based on the previous results (as outlined below). Participants were asked to continue to cancel audible sound at the cancellation cochlea at least every 200 Hz up to 5 kHz. Once cancellation had been attempted from 3-5 kHz, participants were asked to cancel frequencies at least every 100 Hz starting at 2.9 kHz down to 2 kHz. From 2 kHz down to 1 kHz, participants attempted a cancelation frequency at least every 60 Hz.

The prediction employed a cubic spline interpolation and extrapolation from the MATLAB® curve fitting toolbox. Interpolation was used to predict the phase and level of cancellation between two or more frequencies that have already been measured. Spline interpolation is a numerical analysis method which fits input data to a piecewise polynomial. It is particularly suitable for data fitting related to the level differences which can fluctuate considerably over a narrow frequency band with a variable number of peaks and troughs. Spline interpolation was used instead of other data fitting methodologies such as via high order polynomials as those would encounter the problem of the Runge's phenomenon (Tolm, 2014) whereby large prediction errors can occur between the known cancellation values. Data fitting via a moving average would also not be appropriate as it would underestimate the cancellation levels at frequencies where signal summation or destructive interference was occurring.

Spline extrapolation was used when higher or lower frequencies than those already completed were attempted. Safety mechanisms were built in so that if the predicted level was above an intensity threshold the algorithm would present the mean level of the closest three frequencies instead of the level predicted via spline extrapolation. This was necessary to prevent very loud tones from being presented if there was an increasing level trend in the previous values.

By employing the outlined prediction techniques, the data collection time could be reduced to approximately 50 min. If the technique described in Mcleod and Culling (2019) had been used, the experiment would have taken approximately 16 hours for each sitting.

Once frequencies had been attempted from 1 to 5 kHz, participants could use the mouse scroller to sweep the frequency and the prediction algorithm would present what it predicted to be the level and phase differences needed for cancellation at one ear for every frequency. Thus, the sound should remain strongly lateralized as the frequency changed. If not, participants then had the opportunity to attempt further frequencies where the tone had been incompletely lateralized. If a frequency had previously been attempted only the most recent level and phase would be used in the prediction algorithm. This gave a method for correcting mistakes by the participant. Participants were told to keep refining the measurements until a sweep from 1 to 5 kHz and back down to 1 kHz sounded strongly lateralized throughout.

B. Results and Discussion

Fig. 1 shows the phase differences necessary to cancel perceived sound at the left and right cochlea in three participants between 1-5 kHz on five separate experimental sittings. FIG. 2 shows the level differences needed for cancellation at the left and right cochlea on the same five experimental sittings.

Within the same participant, there are similar patterns of phase progression on different sittings with upward and downward inflections of the curve often occurring at the same frequencies. In addition to this, there are pronounced reductions in the level necessary for cancellation over narrow frequency bands. This is most pronounced on the right side in participant 1 at 3.2 kHz and on the left side at 2 kHz in participant 2. A reduction is also visible on the left side in participant 3 at 1.7 kHz. During each instance there is often an associated event in the phase progression where the phase decreases by 180° before resuming the previous phase progression rate. For instance, Participant 1's right-sided cancellation results showed a phase change of 180° between 3.2-3.4 kHz. Sudden phase changes can occur when two signals destructively interfere, leaving a very small resultant. In this case, the phase progression from both BTs was different (as shown in Mcleod and Culling, 2017) and must have been caused by destructive interference, not only at the cancellation cochlea but also at the contralateral cochlea. This is supported by fact that there is a corresponding reduction in cancellation level over the same part of the frequency spectrum. This is an example of an ill condition where crosstalk cancellation would not be successful at this frequency.

There was greater test-retest variability at high and low frequencies when compared to mid (2-4 kHz) frequencies. All participants' phase progression was non-monotonic between 1 and 1.5 kHz, as was previously shown in Mcleod and Culling (2017). Overall the pattern of phase velocity identified in Fig. 1 was very similar to those seen by Tonndorf and Jahn (1981) and Zwislocki (1953).

The pattern of both phase and level variation with frequency is very different between left and right sides in the same participant. As was found in previous studies, there was great variation between sides as well as between participants (Håkansson, et al., 1986; Håkansson et al., 1993; Khalil et al., 1979; Mcleod and Culling, 2017; Stenfelt and Goode, 2005). The fact that the pattern is reproducible across sessions, but the absolute levels are not, was exploited in Exps. 2 and 3. A participant's idiosyncratic pattern of bone conduction was used to rapidly predict a complete transfer function from a small quantity of data at the beginning of a new experimental session

Prior to experimentation, it was anticipated that phase progression with frequency will likely be approximately the same between the left and right side. This is seen in participant 2 and 3 where phase progression between 2.5-4.5 kHz was approximately 370° in both ears of participant 2 and 550° in both ears for participant 3. However, Participant 1's phase progression was 560° for left cancellation and 400° for right cancellation. This discrepancy may have been due to the 180° phase inversion already discussed.

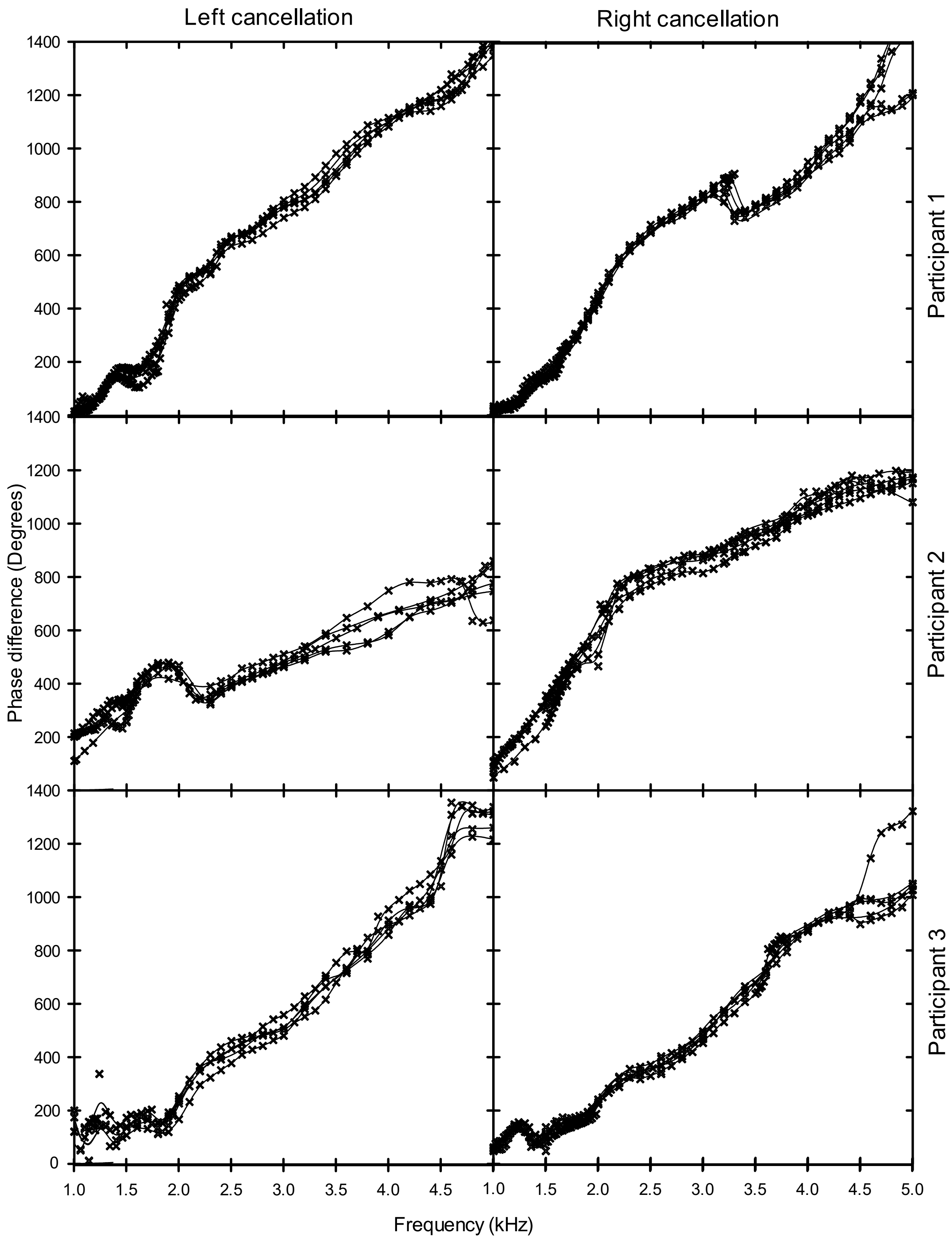

FIG 1. The phase difference needed between bilaterally placed bone transducers to cancel perceived sound at the left and right cochlea on 5 different sittings in three different participants. Line of best fit created using spline fitting method (See procedure).

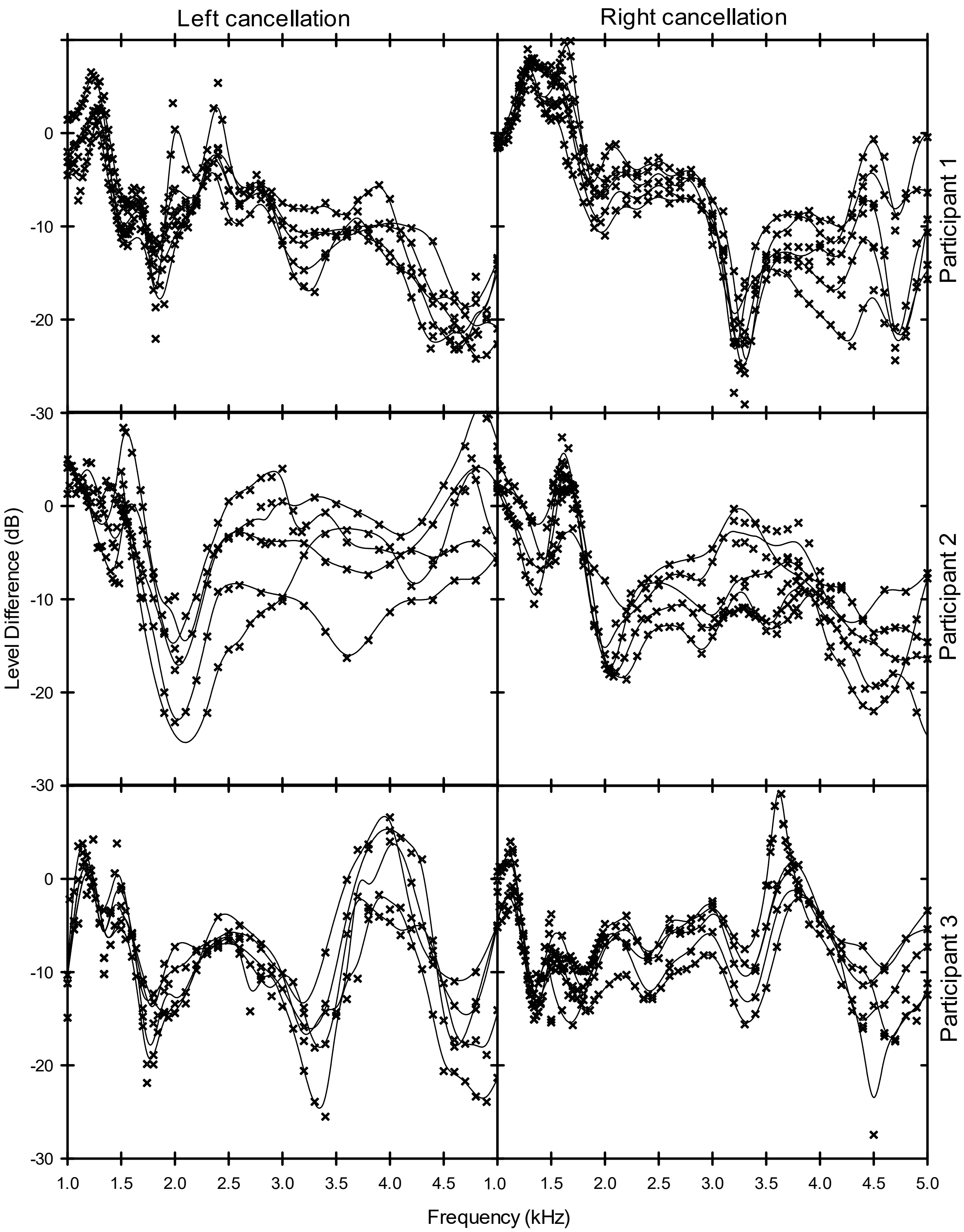

FIG. 2. The level difference needed between bilaterally placed bone transducers to cancel perceived sound at the left and right cochlea on 5 different sittings in three different participants. Line of best fit created using spline fitting method (see Procedure).

III. Experiment 2: Tone reception thresholds

Exp. 2 implemented unilateral crosstalk cancellation and evaluated its effectiveness by measuring masked thresholds for pure tones. Phase and amplitude measurements for cancellation at each ear were made first using methods similar to those from Exp. 1 and then tone reception thresholds were measured with and without a cancellation noise derived from those measurements.

A. Methods

1. Equipment

The same equipment was used as in Experiment 1.

2. Stimuli

Speech-shaped noise maskers were made by filtering Gaussian noise with a 512-point finite impulse response which was matched to the long-term excitation pattern of speech (Lavandier and Culling, 2010; Moore and Glasberg, 1983). The 4-second length of noise was then band-pass filtered using a second 512-point filter to match the frequency over which cancellation had been performed (1-5 kHz). In the noise-only condition (without crosstalk cancellation), twenty individual monaural noise recordings were prepared and used at random in the threshold task.

To create the cancellation noise, the interferers were converted into the frequency domain to obtain the phase and level components. The phase and level differences from the two-BT cancellation task (which the participant had just completed) were then used to alter the level and phase to produce a stimulus whose amplitude at the cochlea would match that of the noise crosstalk and whose phase would be the inverse. Eqs. 1 and 2 from Mcleod & Culling (2019) were used for calculating the crosstalk cancellation signal. The new 'cancellation noise' was then produced by inverse Fourier transform so that it could be added to the tone stimulus. Twenty such paired noise and cancellation noise samples were prepared and used at random in the threshold task described below.

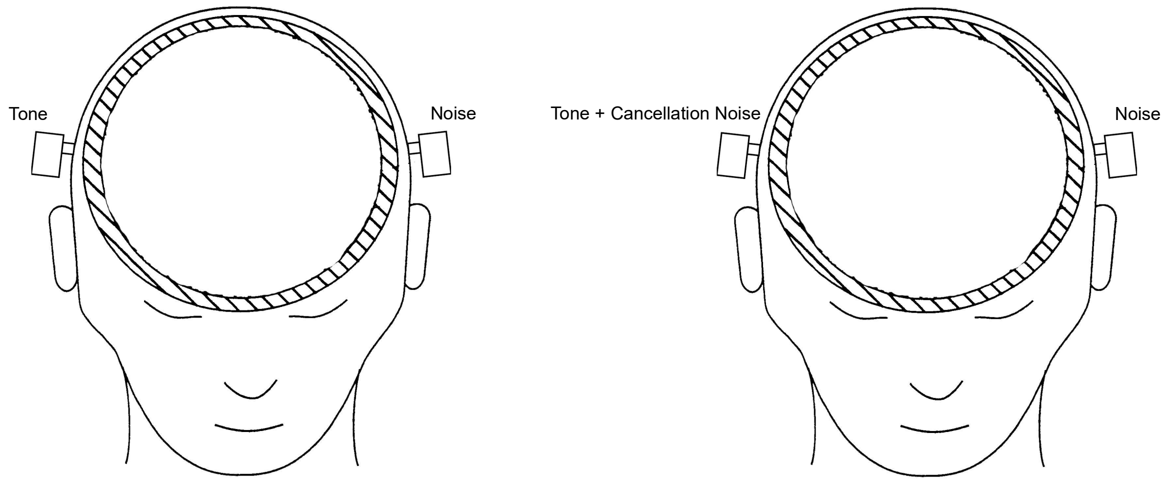

FIG 3. The two main conditions of Exp. 2: a) shows pure tone on one BT and noise on the contralateral BT; b) shows the addition of cancellation noise at the BT with the tone.

3. Participants

The same three listeners participated as in Exp. 1.

4. Procedure

In order to further increase the speed of phase and level data collection a different data prediction algorithm was used prior to masked threshold testing. This was necessary due to the discomfort of wearing a relatively tight headband for a long period of time. The prediction algorithm increased the speed of the measurement by first setting the phase and level parameters as close as possible to the correct values at the beginning of the measurement, thereby reducing the time for the participant to explore the search space. The mean phase and level were measured in the same way as in Exp. 1 every 20 Hz between 1-5 kHz. The participant would attempt cancellation using initial phases and levels that were predicted from their results in Exp. 1. Adjustments to the phase and level differences between the two BTs could then be made via the use of a mouse scroller to refine these parameters. When the participant moved to a new frequency, the measurement speed was further facilitated by combining the mean phase and level results for cancellation from Exp. 1 with the new data to determine the next predicted phase and level. For example, if the participant attempted 3 kHz and found the phase difference to be 20° and the mean change between 3 kHz and 3.1 kHz from Exp. 1 was 30° then the computer would present a phase difference of 50° at 3.1 kHz. This could then be refined by the participant using the same mouse scroller method. If no sound was perceived at the cancellation cochlea, the participant could further adjust the frequency, searching for regions of imperfect cancellation.

Each participant performed 12 runs of detection thresholds (two conditions at six frequencies) which lasted approximately 45 minutes. In order to assess how effective crosstalk cancellation can be at different frequencies, pure tones were tested approximately every 2 equivalent rectangular bandwidths (Moore and Glasberg, 1983) between frequencies 1 and 5 kHz. The test frequencies were 1200, 1530, 1945, 2475, 3150 and 4035 Hz.

Each run utilized a 2-down/1-up adaptive threshold measurement task (Levitt, 1971), with 12 reversals. A 4-dB step size was used for the initial two reversals and 2 dB in subsequent reversals. The average signal level from the last eight reversals was recorded as the threshold level. Each trial consisted of a two-interval, forced-choice task. Each interval lasted 2 seconds with a 0.5-second inter-stimulus interval. The target tone was 0.5 seconds duration and centered within one of the intervals. The participant indicated via button press on a computer terminal which interval contained the target tone. Intervals with and without a target tone were presented in a random order and trial-by-trial feedback was given. The conditions (as shown in Fig. 2) as well as the order of frequencies attempted were randomized to minimize practice effects.

B. Results and Discussion

Fig. 4 shows the mean tone reception threshold (TRT) with and without crosstalk cancellation. A repeated-measures ANOVA was conducted across the two conditions (with/without crosstalk cancellation) 6 frequencies and 3 participants, using the 3 repeat measurements as the random factor. There was a significant improvement in mean thresholds with the addition of cancellation noise [$F(1,2)=515$, $p<0.005$] and a significant reduction in thresholds with increasing tone frequency ($F(5,10)=4.3$, $p<0.05$), which is consistent with the use of speech-shaped noise. No other effects or interactions were significant.

Participants 1, 2 and 3 had similar reductions in TRT with the addition of crosstalk noise. Averaged across frequency, they showed benefits of 11.2 dB, 13 dB and 12.1 dB, respectively. The smallest mean gain in TRT was at the lowest test frequency of 1200 Hz where a 9.2 dB improvement in TRT was identified with addition of crosstalk noise. The frequency with the greatest benefit in TRT with crosstalk noise was at 2475 Hz with a 14.1 dB benefit.

TRTs were collected at six different frequencies in order to more fully assess how accurately the required phase and level differences had been measured across frequency range, as well as to give an indication of the possible benefits of crosstalk cancellation at different frequencies. Crosstalk cancellation was only performed on a single side. Although it would have been possible to construct a bilateral crosstalk cancellation method, this would have meant additional target signal at the contralateral BT. This additional target signal would make evaluation of how well crosstalk cancellation was working less clear; in the adopted design the only change is addition of more noise, making it unambiguous that improvements in threshold are caused by cancellation of the noise. It is likely that the differences in results are due to the accuracy of the phase and level measurements across frequency. Mcleod & Culling (2017) found that the subjective quality of cancellation was lower at lower frequencies. Within the present task, the smaller TRT at lower frequencies supports the participants' subjectively reported difficulty of performing the two-BT cancellation task over this frequency range.

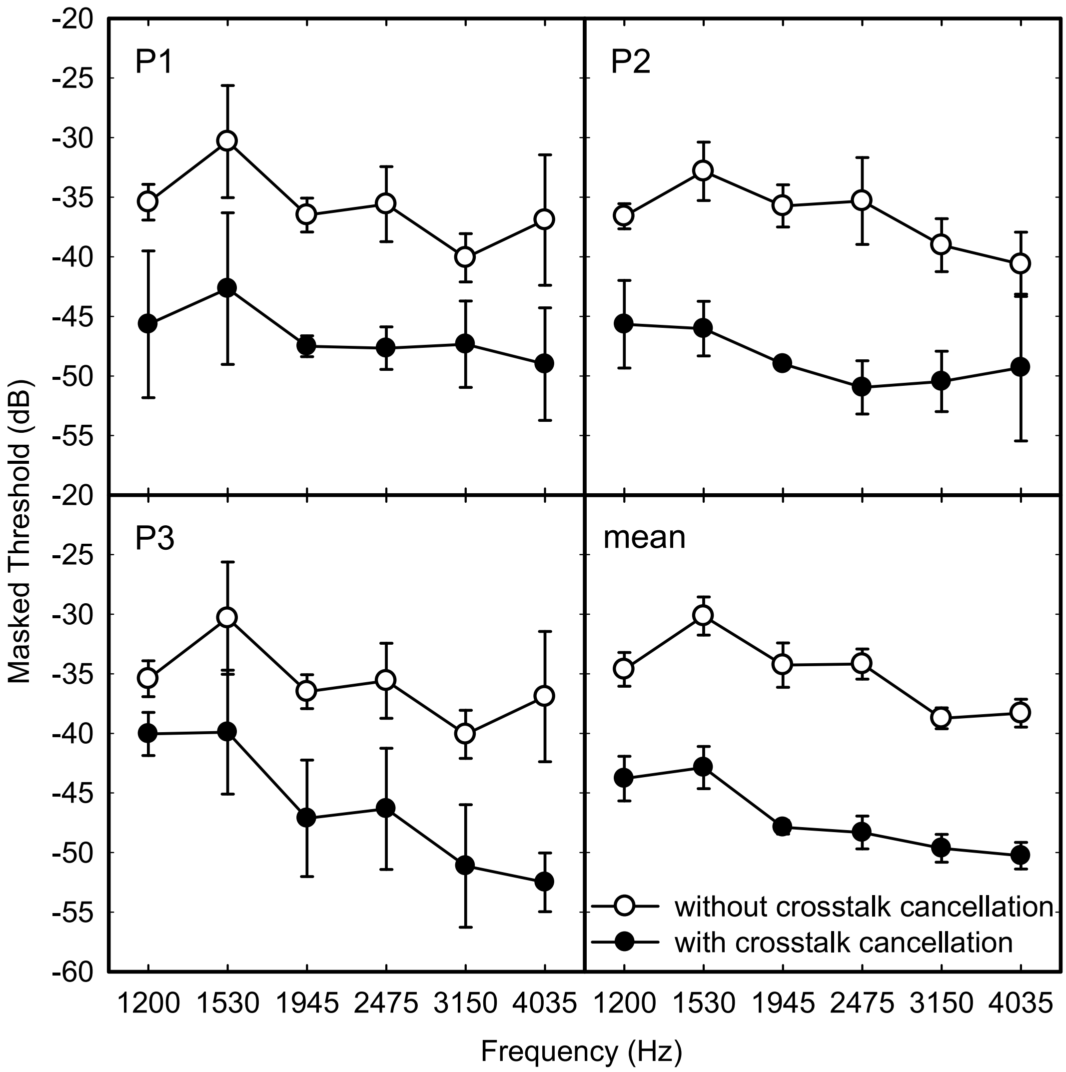

FIG. 4. Tone reception threshold with and without crosstalk cancellation in three participants (3 thresholds per condition) error bars show one standard deviation of the mean.

IV. Experiment 3: speech reception thresholds.

Exp. 3 was similar in structure to Exp. 2. The phase and amplitude values were remeasured and used to implement crosstalk cancellation, but the effectiveness of crosstalk cancellation was then measured through speech reception thresholds (SRTs) with and without cancellation noise.

A. Methods

1. Equipment

The same equipment was used as in Exps. 1 and 2.

2. Stimuli

Speech shaped noise which was then band limited to the range of frequencies over which cancellation data were available (1-5 kHz) was produced using the same method as for the TRTs in Exp 2. Twenty individual monaural noise samples were prepared and used at random in the threshold task. Similarly, twenty stereo noise samples were made with noise on one channel and cancellation noise on the other channel.

Target speech was from a male voice (“CW”) from MIT recordings of the Harvard sentence list (Rothauser et al., 1969). The target speech sentences were also band limited to 1-5 kHz.

3. Participants

The same three listeners participated as in Exps. 1 and 2.

4. Procedure

In each of two experimental sessions, phase and amplitude measurements were initially made using the same method as Exp. 2. These measurements were followed in each session by ten SRTs, five with and five without crosstalk cancellation, producing a total of ten SRTs in each condition for each listener.

A modified version of Plomp's (1986) 1-up/1-down adaptive threshold task was undertaken to obtain SRTs using ten sentences to test each condition. Semantically unpredictable sentences were employed. For example one sentence was “PLUCK the BRIGHT ROSE WITHOUT LEAVES” where keywords are highlighted in capitals (Rothuser et al., 1969). Different sentence lists were employed for each SRT. The procedure aimed to ascertain the signal-to-noise ratio (SNR) where there is 50% intelligibility of the keywords.

The listeners contributed five SRTs for each condition in each of two sessions of approximately 60 mins. At the start of each SRT measurement, the initial SNR for the first target sentence was very low. Participants were instructed to press the "return' key on the keyboard to repeat this stimulus, each time at a 4-dB-higher SNR, until they judged that they could hear two or more target words from the first sentence. They would enter the proposed transcript into the computer program via the keyboard. If one or more of the reported target words matched the target, then the program would display the target sentence on the screen, and participant would self-mark the transcript before moving on to the next target sentence. Otherwise, the first target sentence would be presented again at a 4 dB more favorable SNR, as though the participant had not attempted a transcript. Once recognition of the first sentence had passed this criterion, the remaining nine sentences were presented only once and each transcript self-marked. The SNR decreased by 2 dB if three or more target words were correctly identified or increased by 2 dB if less than three were identified. The average level from the last eight SNRs was used to evaluate the SRT for that condition. The typed transcriptions with self-scoring results were both recorded and visible live to the experimenter in order to verify that the participant complied with instructions.

B. Results and Discussion

FIG. 6 shows the mean SRTs with and without the use of crosstalk cancellation. An ANOVA was conducted across the 3 participants and two conditions (with/without crosstalk cancellation). Ten repeated SRT measurements were taken as the random factor. The crosstalk cancellation produced a significant improvement in thresholds overall of 13.67 dB ($F(1,20)=570$, p<0.001). There were also significant differences between the participants ($F(2,20)=4.13$, $p<0.05$), but no interaction.

In the artificial situation used here, where noise is directed only to one BT and speech to the other, Exp. 3 shows that there can be very large benefits with the addition of crosstalk cancellation noise. However, there are several limitations to the study. Firstly, noise and speech in a real-life scenario are very rarely completely separated at the receivers. It is therefore difficult to show how much of the changes in SRT can be transferred to a real-world scenario. In addition to this, the speech was band limited to cover the same frequency spectrum as the crosstalk cancellation measurements. Thus, our results overestimate any real potential benefits but show that the outlined methodology can be used to create a working crosstalk cancellation system.

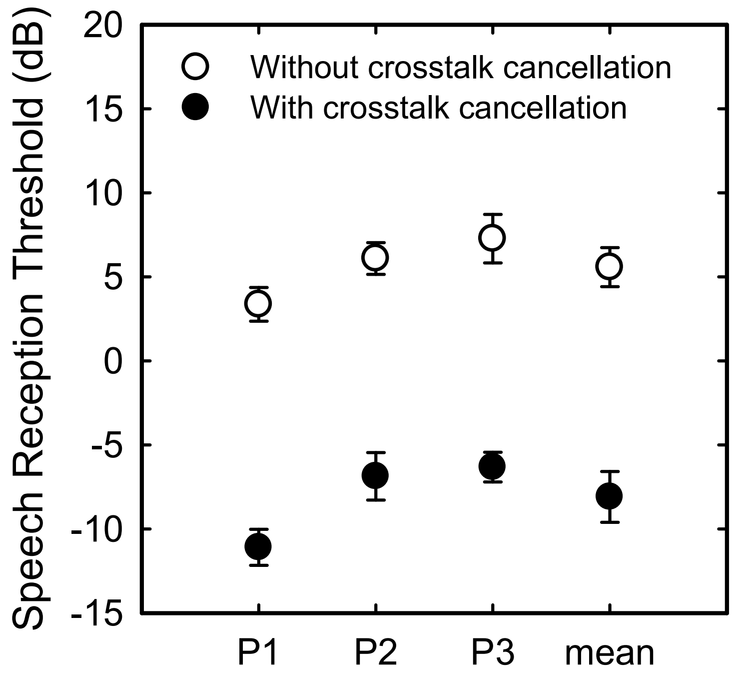

FIG. 5 Mean SRTs with (closed symbols) and without (open symbols) the use of crosstalk cancellation in three participants. Error bars are one standard error of the mean from the sample of 3 repeats for each participant and of 3 participant means for the overall mean.

V. General Discussion

The results presented here have shown that it is possible to psychophysically measure phase and level differences at the cochleae from different bone-conduction sources and that these values can be successfully used in a fixed filter to create a crosstalk cancellation system. The success of the system was evaluated through measuring the masked thresholds at one ear with and without cancellation for both tones and speech. In either case, an improvement in SNR of 10 dB or more was observed.

In order to implement the crosstalk cancellation system in a patient with BCHAs, it would be necessary to feed the microphone signals from each one to the opposite BCHA. Since the phase and level differences from each BCHA are quite different, these signals would need to be filtered with a unique digital filter for each BCHA based on prior psychophysical measurements in the individual patient, and then mixed with the signal from the ipsilateral microphone. It is unlikely that generic filters would be effective, because the transfer function from the abutment to the cochlear will depend on the exact positioning of the abutment, the patient's skull dimensions and any idiosyncratic skull formations that may be associated with their hearing pathology. For users of BCHAs, the fact that the BCHA is coupled to the skull by a permanent titanium abutment should mean that day-to-day changes in coupling, and thus the required filtering are likely to be insignificant. It is, therefore, hoped that that retuning of the filters will be required only occasionally, if at all. Moreover, the current work made very detailed measurements in order to support a demonstration of efficacy. It is likely this methodological rigor could be relaxed to some extent while still obtaining effective crosstalk cancellation. Since the system is intended to unmix the crosstalk occurring within the skull, it will improve stereo separation at the cochlea to something more like that detected at the microphones, regardless of the spatial configuration of sounds externally.

It would be desirable to deliver signals to the two cochleae that were identical to those that would normally be received from airborne sound. The system falls short of this ideal in two ways.

The measurements were limited by practical difficulties to frequencies at or above 1 kHz. The psychophysical task was to detect when one cochlea received little or no stimulation, a situation that can be detected by the listener as a strong lateralization based on inter-cochlear level differences. At lower frequencies, the sound lateralization task was probably disrupted by the listeners' sensitivity to inter-cochlear phase differences. The latter sensitivity normally supports detection of interaural time differences in sound localization. It is limited, for tones, to these low frequencies, but at these frequencies it is thought to be the dominant cue (Wightman and Kistler 1992). Since any adjustment to either the phases or levels delivered by the two bone transducers would affect both the level and phase differences at the cochleae, listeners were faced with a task where they could not isolate and adjust just one cue. Due to this limitation, subsequent tests of masked thresholds were band-limited to the range over which measurements had been possible. As discussed in the Introduction, however, it would, in any case, be unrealistic to implement crosstalk cancellation at low frequencies due to the similarity of the phase at the two cochleae.

The measurements record both the interaural level and phase differences between the bone-conducted sound from the two bone vibrators. In principle, one might hope that this could be used to restore the level and phase differences that would normally reach the cochlea from airborne sound. However, our system concentrated only on restoring the level differences. We took this approach because listeners are relatively insensitive to inter-cochlear phase differences at most of the frequencies that we were able to measure, so the benefits of reproducing the correct phase differences are doubtful. However, there is some sensitivity at high frequencies to envelope delays. It is possible that these survive the effects of phase distortions to some extent, because they are, in effect, short-term level differences. Restoration of sensitivity to high-frequency interaural time delays is thus as possibility with the current approach, but the dominant low frequency interaural time delays cannot be restored.

Since restoration of stereo separation is limited to high frequency level differences, the main likely benefit of the system is the sort of task tested here, the detection of sounds in noise. Spatial release from masking is often dominated by improvement in SNR at one ear or the other (Bronkhorst & Plomp, 1988), and these improvements would be partially obscured by the crosstalk (Stenfelt & Zeitooni, 2013). Unlike sound localization, spatial release from masking is generally unaffected by conflicting cues and seems instead to add together benefits from independent cues and across independent frequency bands (Edmonds and Culling, 2005a,b). The system should thus improve the efficiency with which patients are able to understand speech in background noise situations, employing their two BCHAs to emulate the benefits of binaural hearing.

Future work needs to focus around several areas. Firstly, if the assumption is made that perfect crosstalk cancellation can be achieved to restore inter-cochlear level differences, how much benefit in SRT can be gained in more realistic listening scenarios and how well can this be predicted by binaural models? Secondly what are the benefits in SRT when performing bilateral crosstalk cancellation over the same frequency range with and without band-pass filtering the speech to match the measurement frequencies? Thirdly, how much benefit does crosstalk cancellation confer to sound localization? Finally, there are further challenges regarding how this method can be implemented in real time, since in the outlined scenario all audio was prepared prior to its use. Future research will focus on the development and testing of a prototype low-latency, bilateral crosstalk cancellation system.

VI. Conclusions

Using unilateral crosstalk cancellation of band limited noise, there was a significant benefit in masked threshold measurements with both tones and speech. Future research should focus on ascertaining the potential practical benefits to patients with bilateral bone-conducting hearing aids, as well as the development of a prototype bilateral crosstalk cancellation system that operates in real time.

References

Bauer, B. B. (**1961**). “Stereophonic Earphones and Binaural Loudspeakers,” J. Audio Eng. Soc., **9**, 148–151.

Bosman, Arjan. Snik, Ad F.M. Van Der Pouw, C. T. . T. M., Bosman, A. J., Snik, a. F. M., van der Pouw, C. T. M., Mylanus, E. a, Cremers, C. W. R. J., Mylanus, E. a, et al. (**2001**). “Audiometric evaluation of bilaterally fitted bone-anchored hearing aids,” Int. Journalk Audiol., **40**, 158–167.

Bronkhorst, A. W., & Plomp, R. (**1988**). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," J. Acoust. Soc. Am., **83**, 1508–1516.

Deas, R. W., Adamson, R. B. a, Curran, L. L., Makki, F. M., Bance, M., and Brown, J. a (**2010**). “Audiometric thresholds measured with single and dual BAHA transducers: The effect of phase inversion,” Int. J. Audiol., **49**, 933–9.

Edmonds, B. A., & Culling, J. F. (2005a). The spatial unmasking of speech: evidence for within-channel processing of interaural time delay. J. Acoust. Soc. Am., 117, 3069–3078.

Edmonds, B. A., & Culling, J. F. (2005b). The role of head-related time and level cues in the unmasking of speech in noise and competing speech. Acta Acustica United with Acustica, 91, 546–553.

Håkansson, B., Brandt, A., Carlsson, P., and Tjellstrom, A. (**1993**). “Resonance frequencies of the human skull in vivo,” Acoust. Soc. Am., **95**, 1474–1481.

Håkansson, B., Carlsson, P., and Tjellström, a (**1986**). “The mechanical point impedance of the human head, with and without skin penetration,” J. Acoust. Soc. Am., **80**, 1065–1075.

Khalil, T. B., Viano, D. C., and Smith, D. L. (**1979**). “Experimental analysis of the vibrational characteristics of the human skull,” J. Sound Vib., **63**, 351–376.

Lavandier, M., and Culling, J. F. (**2010**). “Prediction of binaural speech intelligibility against noise in rooms,” J. Acoust. Soc. Am., **127**, 387–99.

Levitt, H. (**1971**). “Transformed up-down methods in psychoacoustics,” J. Acoust. Soc. Am., **49**, Suppl 2:467+.

Mcleod, R. W. J., and Culling, J. F. (**2017**). “Measurements of inter-cochlear level and phase differences of bone-conducted sound,” J. Acoust. Soc. Am., **141**, 3421–3429.

Mcleod, R. W. J., and Culling, J. F. (**2019**). “Psychoacoustic measurement of phase and level for cross-talk cancellation using bilateral bone transducers : Comparison of methods,” J. Acoust. Soc. Am., **146**, 3295–3301.

Moore, B. C. J., and Glasberg, B. R. (**1983**). “Suggested formulae for calculating auditory-filter bandwidths and excitation patterns,” J. Acoust. Soc. Am., **74**, 750–753.

Plomp, R. (**1986**). “A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired,” J. Speech Hear. Res., **29**, 146–154.

Priwin, C., Stenfelt, S., Granström, G., Tjellström, A., and Håkansson, B. (**2004**). “Bilateral bone-anchored hearing aids (BAHAs): an audiometric evaluation," Laryngoscope, **114**, 77–84.

Rothauser, E., Chapman, W., and Guttman, N. (**1969**). "IEEE recommended practice for speech quality measurements," IEEE Trans. Audio Electroacoust.,.

Schroeder, M., and Atal, B. (**1963**). "Computer simulation of sound transmission in rooms," Proc. IEEE,.

Silbiger, H. R., and Sullivan, J. L. (**1969**). "IEEE Recommended Practice for Speech Quality Measurements," IEEE Trans. Audio Electroacoust., **17**, 225–246.

Stenfelt, S., and Goode, R. L. (**2005**). "Transmission properties of bone conducted sound: Measurements in cadaver heads," J. Acoust. Soc. Am., **118**, 2373.

Stenfelt, S., & Zeitooni, M. (**2013**). Binaural hearing ability with mastoid applied bilateral bone conduction stimulation in normal hearing subjects. J. Acoust. Soc. Am., **134**, 481–493.

Tolm, C. D. (**2014**). *Runge ' s phenomenon*,.

Tonndorf, J., and Jahn, A. F. (**1981**). "Velocity of propagation of bone-conducted sound in a human head," J. Acoust. Soc. Am., **70**, 1294–1297.

Wightman, F. L., & Kistler, D. J. (**1992**). "The dominant role of low-frequency interaural time differences in sound localization." J. Acoust. Soc. Am., **9**1, 1648–1661.

Zwislocki, J. (**1953**). "Acoustic Attenuation between the Ears," J. Acoust. Soc. Am., **25**, 752–759.