Intelligibility of bone-conducted speech at different locations compared to air-conducted speech

Raymond M. Stanley and Bruce N. Walker
Sonification Lab, School of Psychology
Georgia Institute of Technology,
Atlanta, GA 30332-0170
bruce.walker@psych.gatech.edu

Bone-conduction transducers offer a unique advantage for radio communication systems, allowing sound transmission while the ear canals remain open for access to environmental sounds, or plugged for blocking of environmental sounds. This study compared the intelligibility of noise-degraded speech presented through bone-conduction hearing administered at different locations, and through air-conduction. Speech intelligibility was assessed using the Diagnostic Rhyme Test. Speech intelligibility was reduced for all of the bone-conduction hearing locations, relative to air-conduction hearing. There were also differences in performance for the various bone conduction locations. These results suggest that given noise-degraded speech, the performance decrement from using bone conduction will have to be weighed against the benefits of being able to dynamically block the ear canal, or leave it open, as situations require. Further, the choice of bone conduction transducer location would need to weigh possible performance differences against the various practical advantages of each location.

INTRODUCTION

Radio communications systems are used widely in modern society for emergency response teams, police, military, and other applications. Many of the environments in which these radio systems are used would benefit from the dynamic ability to plug the ears for protection from loud environmental sounds or, when necessary, leave the ears open for access to ambient sounds. For example, a soldier may need to block his hearing during artillery fire, and leave his ears uncovered at other times to aid in situational awareness. Bone-conduction transducers offer this unique advantage over standard air-conduction headphones that cover the ears and are not compatible with the insertion of earplugs. Before a bone-conducted radio communications system can be implemented, however, the ability to understand speech delivered through bone-conduction transducers (BCTs) needs to be thoroughly understood. The present study considers the intelligibility of speech delivered through BCTs at different locations on the head, compared to performance through air-conduction headphones.

In discussing the mounting locations of bone-conduction transducers, the anatomical terms “condyle”, “mastoid”, and “vertex” will be used. Photographs of these mounting locations can be found in Figure 1.

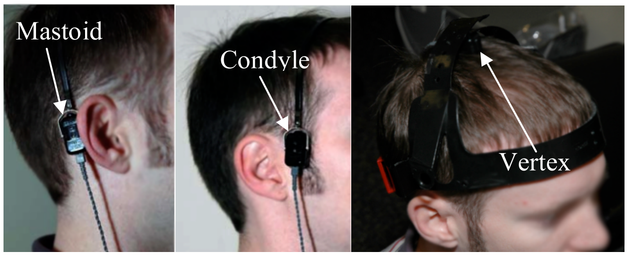

Figure 1. *Mastoid, condyle, and vertex BCT locations.*

Previous Work

Bone-conduction and air-conduction hearing share the same end organ: the cochlea (Békésy, 1960). The difference between the two hearing types lies in the pathway through which the sound travels, and in the way that the cochlea receives the acoustic energy. The three primary mechanisms of bone conduction are: compression of the cochlea via acoustic waves transmitted through the head (Békésy, 1960; Tonndorf, 1966); vibrations of the head interacting with the inertia of the ossicles (Bárány, 1938; Stenfelt, 2006; Tonndorf, 1966); and energy radiated from the head into the ear canal which then follows the air-conduction pathways (Tonndorf, 1966). Most research in bone-conduction hearing has been done within the audiology discipline, due to the diagnostic utility of bone-conduction hearing tests. Specifically, sensitivity to bone-conduction tones is often assessed after air-conduction hearing tests are completed. This allows differentiation of conductive and sensory-neural hearing loss (Robinette & Cevette, 2002).

Research related to the use of bone-conduction transducers for non-clinical purposes in auditory displays has gained recent momentum. Some researchers have begun to investigate using bone-conduction transducers for radio communications. For example, researchers have considered the use of bone-conduction microphones for speech communication in a military context (Acker-Mills, Houtsma, & Ahroon, 2005; Acker-Mills, Houtsma, & Ahroon, 2006). As another example, Walker and colleagues have shown that the intelligibility of multi-talker speech can be improved by spatial segregation of talkers delivered through BCTs (Walker, Stanley, Iyer, Simpson, & Brungart, 2005).

Others have focused on the perception of a single speech channel delivered through BCTs as compared to speech delivered through air-conduction headphones (Gripper, McBride, Osafo-Yeboah, & Jiang, 2007; Osafo-Yeboah, Jiang, Gripper, & Lyons, 2006). Using the Callsign Acquisition Test (CAT), Osafo-Yeboah and colleagues (2006) found no statistically significant difference between speech delivered through headphones and BCTs placed on the condyle. The CAT addresses concerns of having a speech intelligibility test that is representative for a military environment. Specifically, it involves identifying callsign made up of a phonetic alphabet letter code (e.g., alpha, bravo) and a number (e.g., 1, 2) (Gripper, et al., 2007). With more participants and multiple signal-to-noise ratios, Gripper and colleagues (2007), however, found that performance on the CAT was lower for bone-conduction at the condyle (35.7%) than for air conduction (45.6%), and that this difference was statistically significant. Despite the lower performance for speech delivered through bone conduction, it is important to note that high levels of noise were present in the stimulus, suggesting that without this noise, performance could be equally good. The noise was in the form of multi-talker babble presented at 6, 9, and 12 dB higher than the signal (Gripper, et al., 2007).

An important choice in implementing BCTs for radio communications is the location of the transducer. Accordingly, some have considered the effect of the placement of the bone-conduction transducer on speech intelligibility. Osafo-Yeboah and colleagues measured speech intelligibility via the CAT with a BCT located at the forehead, temple, condyle, mastoid, and chin (Osafo-Yeboah, Gripper, McBride, & Jiang, 2006). They found no statistically significant differences in performance between locations. Later work by Osafo-Yeboah and colleagues using the CAT with more participants and other factors effecting intelligibility still did not show a statistically significant effect of location on intelligibility, although the mastoid had slightly higher performance (Osafo-Yeboah, Jiang, McBride, Mountjoy, & Park, 2009). Other work has shown that the lowest threshold for perceiving bone-conducted sounds was found at the condyle, as compared to other locations (McBride, Letowski, & Tran, 2008), which has motivated other work placing the BCT at the condyle (Gripper, et al., 2007; Osafo-Yeboah, Jiang, et al., 2006).

Although the aforementioned work does provide some insight into using bone conduction for radio communications, there needs to be more testing about the effect of BCT location on speech intelligibility—and its performance cost relative to air conduction–to be thoroughly evaluated as a viable means of radio communications. More testing needs to be done with other locations, other speech intelligibility tests, and more participants. After the performance cost of different bone-conduction locations relative to air conduction is better understood, the relative costs and benefits of using such a device can be more thoroughly assessed. The determination of an optimal BCT location can influence how BCTs are designed and incorporated into other gear that rescue and tactical users have to wear.

The Present Study

The purpose of our study was to assess speech intelligibility through bone conduction at the vertex, condyle, and mastoid locations, as well as through standard air-conduction headphones. We have also collected data with participants assigned to conditions with less noise degradation, which showed all conditions at nearly 100% correct. This report will focus on the noise condition in which participants were not at ceiling performance. We chose to assess speech intelligibility with the Diagnostic Rhyme Test (DRT), rather than the CAT that was used in similar previous investigations, and other speech intelligibility tests such as the Modified Rhyme Test (MRT). The DRT provides the strengths that it is efficient enough to allow many within-subjects comparisons within one session; has reliability and validity that has been established through years of standardized use (Voiers, 1983); it has the potential to describe the ability of participants to detect specific types of phonemic features; and avoids use of military callsigns with college students who are not familiar with the vocabulary. Most importantly, assessing speech intelligibility with another test helps identify the robustness of effects (or lack thereof) that previous research has found. Administering the DRT involves having a listener discriminate a pair of rhyming monosyllabic words that differ in their initial consonant.

METHOD

Participants

Seventeen Georgia Tech undergraduates participated in this experiment for course credit. All participants were screened for normal hearing (i.e., they could detect a 20 dB HL pure tone delivered at octaves between 250 Hz to 8000 Hz), and were native speakers of American English. Ten additional participants were excluded due to failure to meet these requirements.

Stimuli & Apparatus

The hearing test was administered using a Micro Audiometrics Corporation DSP Pure Tone Audiometer with TDH headphones. The set of 192 words used to administer the DRT was obtained from Fort Rucker's USAARL lab; a detailed description of the development of these stimuli can be found in Acker-Mills, Houtsma, and Ahroon (2005). These stimuli were created in a reverberant chamber with a background noise intensity of 106 dBA simulating a Black Hawk helicopter in flight (Acker-Mills, et al., 2005). Sixty decibels (A-weighted) of pink noise was also delivered in the room where the participant completed the task, to make the task difficult enough to allow differences between conditions (avoid ceiling effects). The noise was played through Klipsch KSB 1.1 speakers, amplified by a Sony STR-DE597 stereo receiver, which received a signal from a Sony DVP-NS575 DVD/CD player. The signal level of the noise was measured using a Bruel & Kjaer 2260 Sound Level Meter.

Two types of acoustic devices were used in this study: a bone-conduction transducer (BCT) and air-conduction headphones. The bone-conducted stimuli were delivered through a single RadioEar B-71 transducer. These transducers were chosen because of their well-known physical characteristics due to standardized use for Audiometry, and their use in previous literature on bone-conduction hearing. The air-conducted stimuli were delivered through a single earphone of Sennheiser HD 465 supra-aural headphones (which have a diffuse field frequency response).

Stimulus control was provided by a Macintosh PowerMac G4 running Matlab with the Psychophysics Toolbox, displaying visual information on an Apple 17" LCD screen. The audio signal was digitally sent to an M-Audio Firewire Audiophile external sound card, where it was converted to analog form. From here, the signal was sent out to a Denon DRA-275R stereo receiver to power the BCT and a Furman HA-6AB headphone amplifier to power the headphones.

Three BCT locations were considered most relevant from a human factors perspective: the mastoid process behind the ear; on the skull in front of the ear, just above the mandibular condyle; and the vertex at the top-center of the head. The mastoid was chosen because of its closeness to the cochlea, lack of muscular interference, and potential for binaural hearing. The condyle was chosen because of its potential for binaural hearing, closeness to the ear canal, and ease of design as compared to a device that mounts on the mastoid. The vertex was chosen because of its ease of implementation in a helmet. The bone-conduction transducer was secured with a headband provided by RadioEar when it was mounted at the mastoid and condyle locations. For the vertex location, the transducer was secured with the headband assembly from a construction helmet (see Figure 1).

The loudness of all three bone-conducted locations and the air-conducted signal were matched to the loudness of a stimulus playing through a loudspeaker by five pilot participants. The speaker was playing the word "bean" from the DRT stimuli set. The matching was done by having pilot participants turn a knob to adjust the volume of the headphones or BCT, which was intermittently playing with the external speaker at a clearly audible set level.

Procedure

Participants completed an informed consent document, followed by the hearing test. Participants completed the experiment in a room padded with acoustic foam. They donned the headphones or BCTs with instruction from the experimenter. A full DRT (96 trials each) was run with the BCT at each of the three mounting locations, as well as for the headphones. The order of word pairs tested was randomized, and re-randomized each time the DRT was administered. Each trial tested a single word pair. Each trial began with a trigger-press from the participant, which displayed both words, one on each side of the screen. Five hundred milliseconds after the words were displayed, one of the words was played through the acoustic apparatus. The participant then indicated which word they heard by clicking on one of the words with a mouse. After choosing the response, feedback was displayed for one second, which concluded the trial. The visual arrangement of the words (which word was on the left, which one was on the right), as well as which word was presented acoustically, was randomly chosen on each trial. Each session took just over one hour.

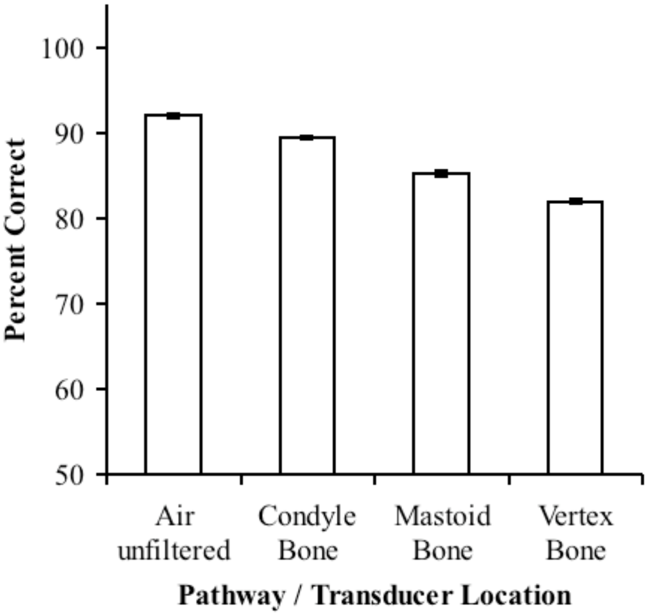

Figure 2. *Total percent correct on the DRT, across voicing features, for each pathway / transducer location. Error bars indicate the standard error of the mean.*

RESULTS

The total percent correct for the DRT in each pathway and transducer location can be seen in Figure 2. The general trend shows air conduction with the best performance, followed by the bone-conducted condyle location, then the mastoid location, and finally the vertex location. Inferential statistics were mostly consistent with this interpretation. Percentage values were converted to proportions then arcsin-transformed to make them suitable for ANOVA treatment (Cohen & Cohen, 1983), and then treated to a one-way repeated measures ANOVA. The ANOVA showed that there was a statistically significant overall effect of pathway / transducer location, $F(3, 48) = 19.926, p <.05$). Post-hoc comparisons executed using Tukey's "Honestly Significant Difference" (HSD) statistic revealed a statistically significant difference between air condition and each of the bone conditions, and a statistically significant difference between the vertex bone and condyle bone condition, but no differences between the other bone-conduction locations (see Table 1). The difference between the mastoid bone and condyle bone approached statistical significance.

Table 1. Post-hoc analyses of percent correct as a function of pathway / transducer location. Critical q(4,48) = 3.78, p < .05. Asterisks indicate statistically significant differences.

Pair	*q*
Air – Bone Condyle	2.76
Air – Bone Mastoid	7.16*
Air – Bone Vertex	9.81*
Bone Condyle – Bone Mastoid	3.72
Bone Condyle – Bone Vertex	8.05*
Bone Mastoid – Bone Vertex	3.26

DISCUSSION

In contrast to previous work finding no statistically significant effect of BCT location on speech intelligibility (Osafo-Yeboah, Gripper, et al., 2006; Osafo-Yeboah, et al., 2009), but consistent with the lowest threshold found at the condyle (McBride, et al., 2008), these data showed that performance for the vertex bone condition was significantly lower than the performance for the condyle bone condition. It is important to keep in mind, of course, that the vertex location was not tested in the previous work. The performance cost associated with the vertex location will have to be weighed against any practical considerations that may make the vertex location more favorable (e.g., the ease of using a BCT that is integrated into a helmet for a tactical or rescue environment).

In addition, the trend of the mastoid bone condition having lower performance than the condyle bone condition approached statistical significance, which differed from the trend shown by Osafo-Yeboah and colleagues (Osafo-Yeboah, et al., 2009), but is consistent with the lowest threshold being found at the condyle (McBride, et al., 2008). There are many possible causes for the difference from Osafo-Yeboah and colleagues (2009), including differences in the intelligibility test used and specific equipment used. Similar studies previously done in our lab using different amplifiers showed a (not statistically significant) trend of better performance at the mastoid location than at the condyle location.

An effect of BCT location is expected, given that differential placement of the transducers creates quite different waveforms in the head (Békésy, 1960). But the effects are not as drastic as one would expect. The discrepancy between physical waveforms and speech intelligibility could be explained by the human auditory system's everyday ability to extract consistent speech features despite considerable variations in the physical signal. The small effect between BCT locations suggests that if a small performance cost can be tolerated, factors besides speech intelligibility can drive decisions about transducer placement. These factors may include effects on transducer coupling due to jaw movements (presumably the mastoid is less effected by talking than the vertex or condyle), integration with other equipment, and user comfort.

These data also replicated the findings of lower speech intelligibility with bone-conduction (Gripper, et al., 2007), with a different test and equipment. In the present study, the lower percent correct on the DRT for all bone conditions, as compared to the air condition, shows that the intelligibility of noise-degraded speech delivered through BCTs was less than when delivered through headphones. This finding is not surprising, given the vastly different pathway that the sounds take for bone-conduction hearing. The performance decrement for bone-conducted speech intelligibility occurring across studies and measures of speech intelligibility indicates that effect is rather robust, and can be expected to occur in a wide variety of situations. However, the fidelity of the devices is likely an important factor to consider, and is discussed below.

Despite its statistical reliability, the magnitude of the degradation in performance does not appear to be terribly large, however – the average performance for all bone-conduction conditions was upwards of about 82 percent, as compared to about 92 percent for air. In addition, recall that the stimuli in this study and others had to be considerably degraded through the addition of noise to avoid ceiling performance in all conditions. Furthermore, compensations for differences or deficiencies in the transducer and the bone-conduction pathway by digital signal processing or filtering (e.g., Stanley & Walker, 2007) could be made to increase the intelligibility of the speech.

The final determination of whether BCTs in general are a suitable alternative to headphones in any given situation will rely on an assessment of the small, but nevertheless present, loss in speech intelligibility for degraded speech, relative to the likely benefit of having the choice of open or plugged ears for access to, or blocking of, environmental sounds. In situations with high-fidelity speech, or in the absence of significant noise, no differences are likely to be seen between air conduction and bone conduction, and thus decisions about pathway can be made entirely based on practical factors.

Limitations & Future Directions

There are several important features to note about this study that must be kept in mind when making decisions about BCT use. First of all, a B-71 BCT was used, which has a steep drop-off above 4 kHz in its frequency response. Although most important features of speech occur within this range, there are still some vocal features (i.e., fricatives) that could be attenuated due to the frequency response. BCTs with better frequency responses are becoming available, and, just like higher quality headphones or loudspeakers, the sound quality should be improved.

In addition, it should be kept in mind that aggregate data are presented here, averaged across participants. Individual users could depart from aggregate trends; in our data, in fact, some participants had trends that differed from the aggregate pattern. Although this could be partially due to chance, there are also individual differences in bone-conduction hearing speech perception that would predict differences.

With information about the optimal location of BCTs for delivering speech communication, it will be important to study exactly how best to design and incorporate BCTs into rescue and tactical equipment and clothing. In addition to location, speech intelligibility could be affected by other factors that should be investigated, including plugging the ears, or wearing a helmet. We also plan to use these results and testing methods to assess the validity of physical models of the bone-conduction pathway developed by collaborators.

CONCLUSIONS

With severely noise-degraded sounds (and with limited-fidelity transducers) there is a performance decrement with bone conduction relative to air conduction. Without this noise degradation, intelligibility of speech sounds presented via bone conduction is comparable to air conduction. Further, different bone conduction locations can lead to different performance levels. In practical applications, such as tactical communications systems, the need for optimal speech intelligibility with a given signal quality will need to be balanced with other specific human factors concerns, such as ear plugs, transducer placement, helmet integration, and so on.

REFERENCES

Acker-Mills, B., Houtsma, A. J. M., & Ahroon, W. (2005). *Speech intelligibility with acoustic and contact microphones.* Paper presented at the New Directions for Improving Audio Effectiveness, Amersfoort, The Netherlands: NATO RTO-MP-HFM-123/RSY, Paper 07.

Acker-Mills, B., Houtsma, A. J. M., & Ahroon, W. (2006). Speech intelligibility in noise using throat and acoustic microphones. *Aviation, Space, and Environmental Medicine, 77*, 26-31.

Bárány, E. (1938). A contribution to the physiology of bone conduction. *Acta Oto-Laryngologica, 26*, 1-223.

Békésy, G. v. (1960). *Experiments In Hearing* (E. G. Wever, Trans.). New York: McGraw-Hill Book Company, Inc.

Cohen, J., & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Second ed.). Hillsdale, NJ: Lawrence Erlbaum.

Gripper, M., McBride, M., Osafo-Yeboah, B., & Jiang, X. (2007). Using the Callsign Acquisition Test (CAT) to compare the speech intelligibility of air versus bone conduction. *International Journal of Industrial Ergonomics, 37*, 631-641.

McBride, M., Letowski, T., & Tran, P. (2008). Bone conduction reception: Head sensitivity mapping. *Ergonomics, 51*(5), 702-718.

Osafo-Yeboah, B., Gripper, M., McBride, M., & Jiang, X. (2006). *Effect of Bone Conduction Vibrator Placement on Speech Intelligibility using the Callsign Acquisition Test – a Pilot Study.* Paper presented at the 2006 Annual Industrial Engineering Management Systems Conference, Cocoa Beach, FL.

Osafo-Yeboah, B., Jiang, X., Gripper, M., & Lyons, L. (2006). *Comparison of speech intelligibility test between air conduction and bone conduction using the callsign acquisition test.* Paper presented at the Human Factors and Ergonomics Society, San Francisco, CA.

Osafo-Yeboah, B., Jiang, X., McBride, M., Mountjoy, D., & Park, E. (2009). Using the Callsign Acquisition Test (CAT) to investigate the impact of background noise, gender, and bone vibrator location on the intelligibility of bone-conducted speech. *International Journal of Industrial Ergonomics, 39*, 246-254.

Robinette, M. S., & Cevette, M. J. (2002). Case history, integrating audiometric results, and clinical decision analysis. In J. Katz (Ed.), *Handbook of Clinical Audiology* (5th ed., pp. 142-156). Philadelphia: Williams & Wilkins.

Stanley, R. M., & Walker, B. N. (2007). Towards a transfer function used to adjust audio for bone-conduction transducers. *Journal of the Acoustical Society of America (Abstract), 123*, 3565.

Stenfelt, S. P. Y. (2006). Middle ear ossicles motion at hearing thresholds with air conduction and bone conduction stimulation. *Journal of the Acoustical Society of America, 119*, 2848.

Tonndorf, J. (1966). Bone conduction: Studies in experimental animals. *Acta Oto-Laryngologica, Suppl. 213*, 1-132.

Voiers, W. D. (1983). Evaluating processed speech using the diagnostic rhyme test. *Speech Technology*, 30-39.

Walker, B. N., & Stanley, R., Iyer, N., Simpson, B. D., & Brungart, D. S. (2005). Evaluation of bone-conduction headsets for use in multitalker communication environments. *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society (HFES2005),* Orlando, FL (26-30 September). pp. 1615-1619.