

Spatial Release From Masking With Bilateral Bone Conduction Stimulation at Mastoid for Normal Hearing Subjects

Trends in Hearing
Volume 28: 1–13
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23312165241234202
journals.sagepub.com/home/tia



Jie Wang¹, Sijia Xie¹, Stefan Stenfelt² , Huali Zhou³ ,
Xiaoya Wang⁴ and Jinqiu Sang⁵ 

Abstract

This study investigates the effect of spatial release from masking (SRM) in bilateral bone conduction (BC) stimulation at the mastoid. Nine adults with normal hearing were tested to determine SRM based on speech recognition thresholds (SRTs) in simulated spatial configurations ranging from 0 to 180 degrees. These configurations were based on nonindividualized head-related transfer functions. The participants were subjected to sound stimulation through either air conduction (AC) via headphones or BC. The results indicated that both the angular separation between the target and the masker, and the modality of sound stimulation, significantly influenced speech recognition performance. As the angular separation between the target and the masker increased up to 150°, both BC and AC SRTs decreased, indicating improved performance. However, performance slightly deteriorated when the angular separation exceeded 150°. For spatial separations less than 75°, BC stimulation provided greater spatial benefits than AC, although this difference was not statistically significant. For separations greater than 75°, AC stimulation offered significantly more spatial benefits than BC. When speech and noise originated from the same side of the head, the “better ear effect” did not significantly contribute to SRM. However, when speech and noise were located on opposite sides of the head, this effect became dominant in SRM.

Keywords

bone conduction, speech intelligibility, speech recognition threshold, spatial release from masking

Received 8 April 2023; Revised received 3 February 2024; accepted 5 February 2024

Introduction

The “cocktail-party problem” (Cherry, 1953) involves frequently encountered listening situations in daily life where the listener focus on a single target sound amidst multiple masking sounds. A key characteristic of such situations is the spatial separation between the target sound and the maskers. As a result of the spatial separation between the target sound and masking sounds, there is a decrease in masking effects and an enhancement in speech recognition performance (Asp et al., 2020; Bronkhorst, 2000; Hawley et al., 1999), commonly known as spatial release from masking (SRM) (Glyde et al., 2013; Hawley et al., 2004; Kidd et al., 2010; Litovsky, 2012). SRM is governed by a multifaceted set of mechanisms, with a pivotal role attributed to binaural hearing. Binaural hearing relies on the temporal and level differences in sound arrival at each ear arising from different spatial locations (Bronkhorst & Plomp, 1988). These differences between the ears, termed spatial

cues, provide information for the auditory system to selectively amplify the signal originating from a specific source while concurrently suppressing competing sources. This results in an improved ability to perceive and comprehend the target auditory signal.

¹School of Electronics and Communication Engineering, Guangzhou University, Guangzhou, China

²Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

³Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen University, Shenzhen, China

⁴Otolaryngology Department, Guangzhou Women and Children's Medical Center, Guangzhou, China

⁵Shanghai Institute of AI for Education, East China Normal University, Shanghai, China

Corresponding author:

Jinqiu Sang, East China Normal University, Shanghai, 200062, China.
Email: sangjinqiu@mail.ioa.ac.cn



The calculation of SRM is usually based on the difference in speech recognition threshold (SRT) between situations where maskers and a target sound are co-located versus cases where they are spatially separated. The target-to-masker ratio (TMR) that produces 50% intelligibility is often used to define the speech recognition threshold. This study aimed to examine the SRM as a function of spatial separation between the target speech and masker in participants with normal hearing. Along with comparing and assessing the SRM produced by two different stimulation modalities (bone conduction and air conduction), specific tasks included determining the minimum angular separation to produce SRM and the angular separation to achieve the maximum SRM with bone conduction (BC) stimulation.

Bone conduction plays a crucial role in auditory perception, including both externally and internally generated sounds (Reinfeldt et al., 2010; Reinfeldt et al., 2007). Analogous to air conduction (AC), BC facilitates auditory perception by stimulating the sensory cells within the cochlea (Stenfelt, 2007; Stenfelt et al., 2003). The fundamental distinction between AC and BC sound transmission resides in their respective mechanisms. While AC primarily relies on the vibration of the middle ear ossicles at the oval window, BC sound transmission is characterized by its complexity, involving a multitude of pathways (Stenfelt, 2011; Stenfelt & Goode, 2005). In a healthy ear, sound perception through BC predominantly results from the vibrations of the inner ear itself (Stenfelt, 2020; Stenfelt & Prodanovic, 2022).

As previously mentioned, in typical auditory perception through AC, sounds exhibit variations in arrival time and level at each ear, dependent on their spatial location. These variations are referred to as interaural time differences (ITD) and interaural level differences (ILD). Established models for assessing SRM using ITDs and ILDs are generally effective in estimating the intelligibility of target speech when both the target and interfering sounds are situated in the horizontal plane (Bronkhorst, 2015).

Recent studies have highlighted that SRM is shaped by a combination of monaural and binaural mechanisms (Arbogast et al., 2002; Best et al., 2006; Freyman et al., 1999; Kidd et al., 1998; Marrone et al., 2008). These mechanisms encompass the phenomena of enhanced listening through the better ear and auditory binaural processing. Based on the head shadow effect, which refers to the attenuation of sound originating from the opposite side due to the listener's head, monaural processing of SRM results in an elevated TMR when the acoustic stimulus reaches the "advantaged" ear. For instance, in scenarios where the target sound is located on one side and the masker sound is on the opposite side relative to the listener, the ipsilateral ear to the target source exhibits a higher TMR, referred to as the "better-ear effect" (Edmonds & Culling, 2006; Hawley et al., 2004).

The spatial separation between the target and masker sounds contributes to binaural processing, which is partially

governed by variations in ITD and ILD (Culling et al., 2004; Glyde et al., 2013; Kidd et al., 2010). The "better-ear effect" becomes less pronounced in many listening scenarios where the target sound is positioned between maskers distributed around the listener, highlighting the pivotal role of binaural processing in SRM.

Recent research has investigated the phenomenon of SRM. Marrone et al. (2008a) discovered that the majority of individuals with normal hearing experienced the full benefits of SRM at 45° spatial separation, with most of the SRM effect concentrated within the initial 15° of spatial separation. A prevalent approach in SRM investigations involves the use of AC headphones and virtual sound environments. Srinivasan et al. (2016) investigated SRM for small spatial separations using AC headphones and reported that young individuals with normal hearing achieved SRM benefits at spatial separations as small as 2° to 4°. Peng & Litovsky (2021) evaluated SRM in children with bilateral cochlear implants (BiCIs) and children with normal hearing utilizing AC headphones. They found that when the target and masking sounds were positioned 180° apart, children with normal hearing exhibited greater improvements in intelligibility compared to previous findings in the literature involving smaller spatial separations (Bronkhorst, 2000, 2015). Furthermore, Peng & Litovsky (2021) quantified the minimal angular separation required between the target and masker to achieve an intelligibility increase from 50% to 70%, defining it as the "minimum angular separation." In that study, over half of the children with BiCIs attained SRM using binaural cues.

In a study by Zeitooni et al. (2016) involving adults with normal hearing, SRM was investigated using both bilateral AC stimulation through earphones and BC stimulation at the mastoid. Their findings suggested that AC stimulation was more effective in providing binaural hearing benefit compared to BC stimulation. Specifically, at a spatial separation of 45°, the SRM achieved through BC stimulation was approximately 50% lower than attained through AC stimulation in terms of decibels.

The studies mentioned above primarily employed AC sound for investigating SRM. Nevertheless, it is worth noting that SRM has also been explored with bilateral BC stimulation (Canale et al., 2022; Denanto et al., 2022; Dun et al., 2013). The fundamental distinction between bilateral AC and BC stimulation lies in their respective transmission characteristics (Stenfelt, 2005). While AC stimulation leads to minimal cross-head transmission (i.e., sound transmission from the ear on one side to the ear on the opposite side), BC stimulation introduces small time and level differences in sound reaching both ears (Stenfelt, 2012; Surendran & Stenfelt, 2023; Zwislocki, 2005). This cross-head transmission with BC is believed to impede SRM and implies that findings with AC cannot be directly extrapolated to BC.

Bilateral application of BC sound is important primarily in two domains: (1) BC hearing aids (BCHAs) (Maier et al.,

2022) and (2) the use of BC headsets in consumer electronics (Surendran et al., 2023). Bilateral application of BCHAs has shown promise in enhancing SRM, despite the inherent drawback of cross-head sound transmission (Colquitt et al., 2011; Heath et al., 2022). However, it is essential to underscore that further investigations are required to understand the advantages and implications of bilateral BCHA application. Conversely, the knowledge remains comparatively limited regarding the use of BC in virtual reality applications. This area has gained attention due to proposed advantages of an open ear canal and the potential for environmental sound awareness (Lindeman et al., 2008; Voong & Oehler, 2019). Nevertheless, a comprehensive understanding of the benefits and limitations of BC in virtual reality applications necessitates further exploration and investigation.

The aim of this present study is to investigate the impact of various spatial separations between speech and masking noise on speech recognition when using bilateral BC stimulation. This investigation will seek to identify the minimum spatial separation necessary to elicit SRM and determine the spatial separation required to attain maximum SRM when employing BC stimulation. Furthermore, the study aims to compare SRM outcomes obtained through BC and those obtained via AC using earphones. These aims were realized through the investigation of the SRT under varying spatial configurations within the frontal horizontal plane. This configuration enables a comprehensive analysis, encompassing the assessment of the better-ear effect and the influences of binaural processing.

Material and Methods

Participants

The required sample size was estimated based on analysis of variance (ANOVA) using G*Power software (Kang, 2021). With an effect size of 0.25, the calculated number of participants needed for the study was seven. Ultimately, a total of nine participants with normal hearing (comprising six males and three females) within the age range of 23 to 28 years were recruited for the study. All selected participants were native speakers of Mandarin and possessed no prior history of otologic disorders. To affirm their normal hearing status, AC pure-tone audiometry was conducted. These assessments covered octave frequencies ranging from 0.25 to 8 kHz. The results of these assessments indicated hearing thresholds of 20 dB HL or better, with no interaural differences exceeding 10 dB at any tested frequency.

Apparatus

The experiment was conducted within a sound proof room, with a background noise of 25 dB sound pressure level (SPL). Computer-generated stimuli were delivered through a sound card (Fireface UFX II, RME, Haimhausen, Germany), connected to BC transducers (B-81, RadioEar,

Middelfart, Denmark) and AC headphones (HD 650, Sennheiser, Wedemark, Germany). During the BC trials, participants underwent testing with bilateral BC transducers positioned at the mastoid secured by an elastic band producing a static force of approximately 3N. The BC transducers were placed once, and all BC measurements were made with them in the same place. Care was taken to avoid the BC transducers to touch the pinnae and to place them as symmetrical as possible on the two mastoids.

Stimuli

Sentence materials from the Mandarin Hearing In Noise Test (MHINT) corpus were used as the target speech (Wong et al., 2007). The MHINT corpus consists of 14 lists, each including 20 sentences with 10 words per sentence (e.g., “The apples in the orchard are big and red”). All lists were used in this study. A babble noise synthesized from 12 randomly selected speech segments in the MHINT corpus was used as the masker, and the babble noise had similar frequency spectrum to the target speech (Kalikow et al., 1977).

The Head-Related Transfer Functions (HRTFs) were convolved with both the target and masker sounds, thus generating simulations with the desired spatial configurations. The HRTFs used in this study were from the MIT database, and they were originally acquired in an anechoic chamber, using a KEMAR manikin and measured at the ear-canal entrance.

A total of 15 spatial configurations in the horizontal plane, as illustrated in Figure 1, were examined. In all cases, the target speech remained positioned at the right side of the listener, which was defined as 0°. The masker was either co-located with the target (i.e., at 0°) or spatially separated from the target across 14 conditions, with the masker’s position varying at 5°, 10°, 15°, 30°, 45°, 60°, 75°, 90°, 105°, 120°, 135°, 150°, 165°, or 180°. For instance, the designation “SON5” represents the condition where the target speech was situated at 0°, and the masking noise was positioned at 5°.

Experimental Procedures

To ensure comparability between outcomes generated by AC and BC, an equalization process was employed to match the perceived stimulation levels of the two modalities. The loudness balancing procedure used the babble noise, which served as masker during SRT measurements. During the procedure, participants were presented alternating with the stimulus via AC at 65 dB SPL (calibrated on a dummy head) and through BC at the ipsilateral ear. Participants were tasked with adjusting the BC stimulus until it subjectively matched the perceived loudness of the AC stimulus (Qin & Usagawa, 2017). The adjusted level of the BC stimuli for each individual was stored and subsequently used in the experiments.

Each participant was tested in 30 conditions, resulting from the combination of two stimulation modalities (AC and BC) and 15 spatial configurations. Within each of

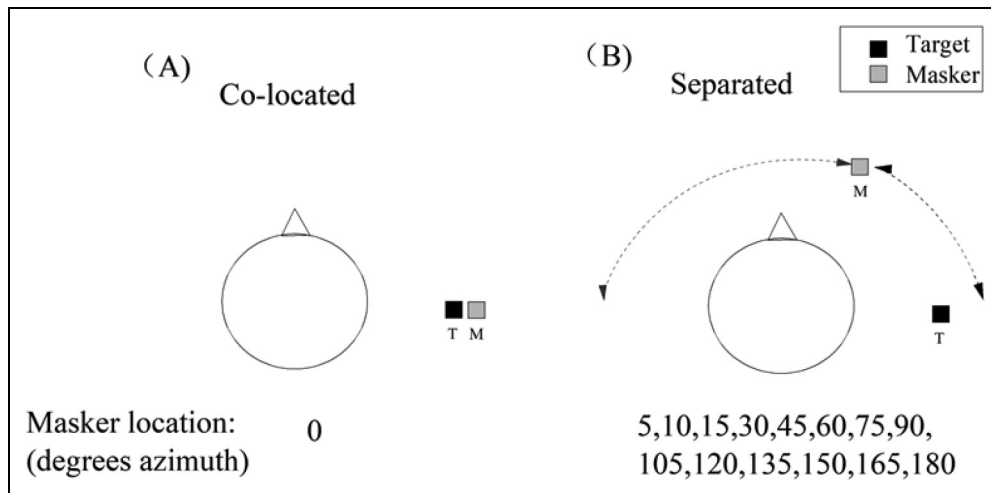


Figure 1. The setup for measuring speech recognition in noise is illustrated. In the co-located condition (A), the target, T, and masker, M, were located at 0° azimuth. In the spatially separated condition (B), the target was located at 0° azimuth, and the masker was located around the target at 5°, 10°, 15°, 30°, 45°, 60°, 75°, 90°, 105°, 120°, 135°, 150°, 165°, or 180° azimuth. The black square denote the target and the gray square denote the masker.

these conditions, an SRT was acquired using one MHINT list. To prevent order bias, the presentation order of both lists and sentences within each list was randomized. In the current study, for practical reasons, all AC testing were done at first session and all BC testing were done at second session. To minimize memory effects of the speech test, the time between the two sessions was 2 months.

The SRT, defined as the TMR yielding 50% intelligibility, was determined through an adaptive one-down-one-up procedure (Levitt, 1971). In each trial, a sentence embedded in noise was presented to the participants. The sentence could be repeated up to three times based on the participant's response. Participants were instructed to orally repeat the target sentence without receiving any feedback regarding the accuracy of their responses. A trial was considered correct if more than five keywords within the sentence were accurately identified; otherwise, it was considered incorrect. Correct identification of a sentence led to a reduction in the level of the subsequent sentence, while maintaining a constant noise level. Conversely, incorrect identification resulted in an increase of the target speech level for the next trial.

Throughout the experiment, the masker level was set at 65 dB SPL, and the TMR was initially set to 5 dB before being adaptively adjusted. The SRT was determined by computing the mean of the TMRs for the last eight sentences, with an initial step size of 8 dB, followed by reductions to 4 dB for the subsequent two reversals, and further reduced to 2 dB after the fourth reversal.

Data Analysis

Following the SRT measurements, SRM was computed as the difference in SRTs observed when the target and masker were spatially separated compared to when they

were co-located. These computations were performed separately for both AC and BC stimulations. To investigate the influence of the stimulation modalities on SRM, repeated measures ANOVA was conducted on the results obtained from AC and BC stimulation. Additionally, a nonlinear polynomial function of spatial separation was fitted to the SRT data for both conduction pathways to further examine the relationships between spatial separation and SRM. To deepen our understanding of the differences between AC and BC stimulation regarding SRM, the magnitude of SRM achieved with AC stimulation was quantified in terms of the spatial separation required to attain a comparable SRM with BC stimulation.

Results

Speech Recognition Thresholds

Figure 2 displays the mean SRTs measured in all 15 conditions based on data from all participants. A two-way repeated-measures ANOVA was conducted on the SRT values with spatial separation and stimulation modality as within-subject factors. Significant effects of spatial condition ($F(14,224) = 440.33, p < .001$) and stimulation modality ($F(1,16) = 102.88, p < .001$) were observed, with a significant interaction between spatial separation and stimulation modality ($F(14,240) = 16.92, p < .001$).

Post hoc comparisons with Bonferroni corrections were conducted to investigate the effect of spatial separation on the BC results with S0N0 as the control condition. Figure 2 shows a trend of decreasing SRT with increasing spatial separation. The analysis indicated no statistically significant difference ($p = .15$) in SRT between the conditions S0N5 (mean -6.8 dB) and S0N0 (mean -4.6 dB). However, at

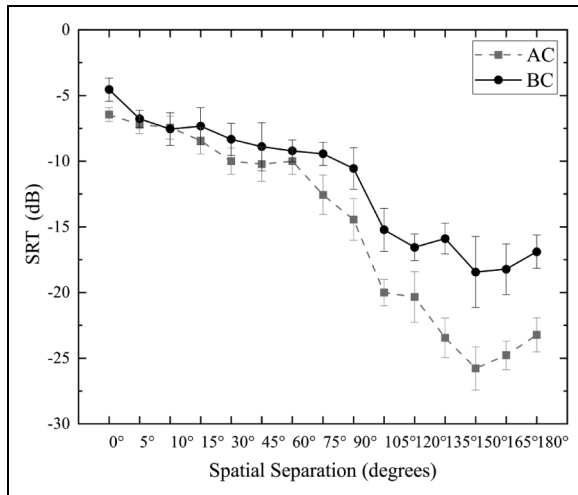


Figure 2. Mean speech recognition thresholds measured at 50% speech intelligibility by BC stimulation and AC stimulation at all 15 spatial conditions. The vertical lines represent the standard error of the mean. Black symbols denote mean SRT with BC while gray symbols denote mean SRT with AC stimulation. AC = air conduction; BC = bone conduction; SRT = speech recognition threshold.

test condition SON10 (mean -7.6 dB) and higher spatial separations, the reduction in SRT was statistically significant ($p < .05$). This indicates that, at separations of 10° or greater, the SRT was significantly better than that observed when speech and masker were co-located. The most favorable SRT, characterized by the lowest value, was observed in condition SON150, with a mean of -18.4 dB. For spatial separations exceeding 150° , the SRT for BC stimulation deteriorated, indicated by an increase in values.

For AC stimulation, a similar trend of decreasing SRT with growing spatial separation was observed as in BC stimulation, as shown in Figure 2. Furthermore, similarly as in BC stimulation, the lowest SRT for AC stimulation was also at SON150, with a mean of -25.8 dB (a 7.4 dB improvement compared to BC stimulation). For spatial separations exceeding 150° , the SRT for AC stimulation also deteriorated, indicated by an increase in SRTs. Post hoc comparisons with Bonferroni corrections were also conducted on the AC results with SON0 as the control condition. In contrast to BC stimuli, significant differences in SRT for AC stimuli were observed only when the spatial separation was 30° or more. Specifically, SRT for SON30 and larger spatial separations were significantly different from SON0 ($p < .05$).

When comparing the SRTs between stimulation modality (BC vs. AC), Figure 2 shows overall higher BC stimulated SRTs compared to AC stimulated SRTs. Post hoc comparisons with Bonferroni corrections were used to investigate the effect of stimulation modality on the SRTs. The results show that the SRT differences between AC and BC stimulation were not statistically significant for spatial separations ranging between 5° and 60° ($p > .05$). For the other spatial

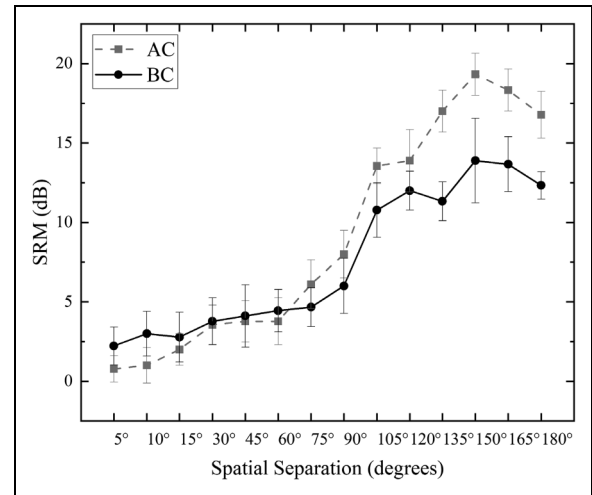


Figure 3. Mean spatial release from masking for BC stimulation and AC stimulation in all 15 spatial conditions. Black and gray symbols denoted mean SRM with BC and AC stimulation, respectively, while the error bars represent the standard error of the mean. AC = air conduction; BC = bone conduction; SRM = spatial release from masking.

separations, the AC stimulated SRTs were significantly lower than those by BC stimulation ($p < .05$).

Spatial Release From Masking

Figure 3 shows the average benefit in speech intelligibility across the 14 conditions with separated speech and noise in comparison to the co-located condition, that is, the SRM. The participants showed nearly a continuous improvement in SRM with separation in the range from 5° to 150° , shown by the positive and increasing SRM values in Figure 3. To assess these SRM values, a repeated-measures ANOVA was conducted, with spatial separation and stimulation modality (AC and BC) as within-subject factors. The analysis revealed significant main effects for both spatial separation ($F(13,208) = 379.42$, $p < .001$) and stimulation modality ($F(1,16) = 16.82$, $p < .001$), as well as a significant interaction between spatial separation and stimulation modality ($F(13,224) = 14.49$, $p < .001$).

At SON5 the SRM with BC stimulation was 2.2 dB which increased to 3.0 dB at SON10. As stated previously, the test condition SON150 gave the highest improvement, which was 13.9 dB for SRM with BC stimulation. The SRM results for AC stimulation showed a comparable pattern but generally higher to those observed with BC stimulation. AC sound also produced the largest SRM under the test condition SON150 with a value of 19.3 dB. Consequently, the spatial benefit with AC stimulation was 5.4 dB higher compared to BC stimulation at SON150.

Post hoc analysis with Bonferroni corrections was conducted to explore the effects of spatial separation on the

SRMs with AC and BC stimulation, as well as the effect of stimulation modality on the SRM for each spatial separation. When the stimulation was by BC, the least separation to produce a statistically significant improvement was 10°. Specifically, the 2.2 dB improvement at SON5 did not reach statistical significance whereas the 3.0 dB improvement at SON10 demonstrated statistical significance. For all other separations, the SRM with BC stimulation was significant.

In the current sample, it required a spatial separation of 30° for the SRM with AC stimulation to reach statistical significance. Consequently, the AC generated SRMs with 5°, 10°, and 15° spatial separations were not statistically significant whereas all spatial separations of 30° or larger were statistically significant. Regardless of whether AC or BC sound was employed, participants achieved the greatest benefit from spatial separation when the noise source was located 150° away from the target speech.

Across all spatial conditions, AC stimulation gave overall 4.4 dB higher SRM compared to BC stimulation. When the spatial separation was <75°, meaning that both the target speech and masker were situated on the same side of the participants' head, SRM with BC transmission was marginally higher than with AC stimulation. However, these differences were statistically significant only at SON5 and SON10 ($p < .05$). Conversely, when spatial separation equaled or exceeded 75°, SRM for AC stimulation was significantly greater than that for BC stimulation ($p < .05$).

Discussion

Equivalence Between Speech Recognition Thresholds for AC and BC

This section focuses on investigating the effect of spatial separations on the speech recognition thresholds for the two stimulation modalities. SRTs with both AC headphones and BC transducers exhibited similar patterns, but the participants achieved lower SRTs in the AC condition compared to the BC condition.

One way to address the difference in SRTs between AC and BC is to compute the spatial difference at which they produce the same SRT. As a first step, the data was fitted to a parametric polynomial function using a nonlinear curve fitting function (Matlab function "poly")

$$SRT = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} \quad (1)$$

where a_0, a_1, \dots, a_{n-1} are the fitted coefficients and x the spatial separation in degrees.

Figure 4 shows the fitted curves of the spatial separation parameters with the SRT data for AC and BC (Pearson correlation coefficient: AC: 0.991; BC: 0.994). In the test condition SON0 with BC stimulation, there are no benefits of stimulation modality or spatial separation. This condition

was used as the baseline to assess the advantages of AC stimulation and spatial separation. The benefit with AC stimulation was determined as the SRT difference between the baseline condition and SON0 with AC stimulation.

The function fitted to the BC SRT data was used to identify the spatial separation that yielded a benefit equal to the benefit of AC stimulation, which was 2.8° (where the fitted SRT function for BC stimulation was -6.4 dB). At this separation, the SRM with BC stimulation was 1.9 dB. Consequently, the benefit of AC stimulation equaled a BC separation of 2.8° at the collocated position. Similarly, the AC function required only 103.6° spatial separation to produce the same SRT (-18.6 dB) as the lowest SRT with the BC function (at 159.8° spatial separation). The lowest value of the function for AC stimulation was a SRT of -25.9 dB at 159.8° spatial separation, a value not reached by BC stimulation. Table 1 indicates the additional spatial separation of a BC sound to reach the same SRT as with AC stimulation as a function of separation between target and masker in the current setup.

As shown in Figure 4 and Table 1, the spatial separation required for BC stimulation was always larger than for AC stimulation to produce the same SRT. For different SRTs, the additional spatial separation required for BC stimulation compared to AC stimulation varied. However, when the target sound and the masker were presented as BC stimuli in the same hemifield, that is, at spatial separations between 0° to 75°, the additional spatial separation required for the BC stimulated SRTs to equal AC stimulated SRTs increased with increasing spatial separations. One possible

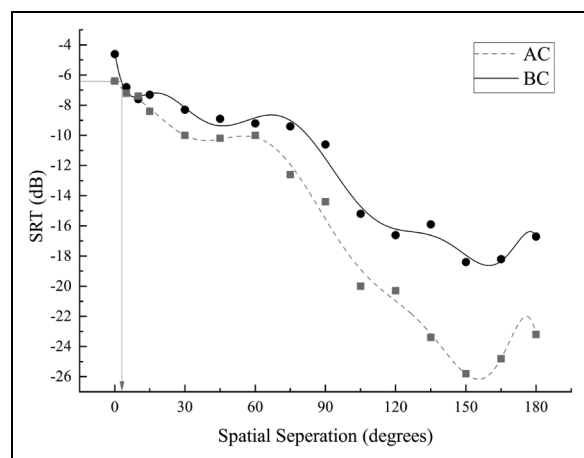


Figure 4. Speech recognition thresholds as a function of spatial separation with BC stimulation and AC stimulation in all 15 spatial conditions. The polynomial functions fitted to the data are represented by the solid and dashed lines. Black and gray symbols denote measured mean SRTs with BC and AC stimulation, respectively. The gray arrows indicate the spatial separation on the BC function that has an SRT comparable to the AC condition at SON0. AC = air conduction; BC = bone conduction; SRT = speech recognition threshold.

Table 1. The Equivalence of AC and BC for all Spatial Conditions.

Location of masker	Fitted SRT for BC	Equal AC performance	Additional spatial separation for BC
0°	-4.5 dB		
5°	-7.1 dB	4.9°	0.1°
10°	-7.4 dB	8.2°	1.8°
15°	-7.2 dB	6.1°	8.9°
30°	-8.2 dB	14.6°	15.4°
45°	-9.4 dB	23.8°	21.2°
60°	-8.8 dB	19.7°	40.3°
75°	-9.0 dB	21.0°	54.0°
90°	-11.5 dB	72.9°	17.1°
105°	-14.7 dB	86.7°	18.3°
120°	-16.2 dB	92.7°	27.3°
135°	-16.6 dB	94.3°	40.7°
150°	-17.9 dB	100.2°	49.8°
165°	-18.3 dB	102.11°	57.98°
180°	-16.7 dB	94.6°	85.4°

AC = air conduction; BC = bone conduction; SRT = speech recognition threshold.

reason for this finding is that binaural cues become more significant with the separation of target and masker in the same hemifield.

One mechanism for the worse BC SRTs compared to AC SRTs with target and masker separation is the BC interaural transmission with stimulation at the mastoid (Surendran & Stenfelt, 2023; Reinfeldt et al., 2013; Wang et al., 2023). Due to the crosstalk between the ears, the binaural cues with BC stimulation were disrupted (McBride et al., 2015). This required a larger spatial separation for BC stimulation relative to AC stimulation as the spatial separation increased. When the target sound and masker with BC stimulation were presented in different hemifields, that is, at spatial separations between 90° and 180°, the required additional spatial separation increased with increasing spatial separations. However, the additional spatial separation required for BC stimulation to equal AC stimulation was smaller at spatial separations between 90° and 150° than at a spatial separation of 75°. This is believed to be caused by the monaural head shadow effect in effect at separations of more than 90°. Even if this leads to improvement with both AC and BC stimulation, part of the better ear effect is already surpassed by the binaural benefit with AC stimulation. The better ear effect caused by the head shadow dominated the BC speech recognition performance when the spatial separation was between 120° and 180° (Stenfelt, 2005).

Effects of Spatial Separation on SRT and SRM

Figure 2 illustrates that the SRT was less favorable (higher) with BC stimulation when compared to AC stimulation. One possible explanation for these differences between AC and BC stimulation lies in the limited binaural isolation

achieved with BC, as opposed to AC stimulation (McBride et al., 2015; Surendran et al., 2023). The existence of interaural crosstalk in BC stimulation can distort binaural cues, thereby contributing to the observed higher SRT (McLeod & Culling, 2020; Stenfelt, 2005).

Zeitouni et al. (2016) measured the SRTs in scenarios where the target and maskers were either co-located from the front or spatially separated, with speech from the front and noise positioned at 45° from the front. Their findings revealed that the SRTs were nearly identical with both stimulation modalities when speech and noise were co-located. However, AC stimulation outperformed BC stimulation by approximately 4 dB for spatial separation of 45° between the speech and noise. This discrepancy was attributed to the presence of BC crosstalk. Another possible explanation for these differences is the bandwidth of the signals. It has been shown that a skin interface for BC stimulation limits the high-frequency content of the signal (Stenfelt & Zeitouni, 2013; Zeitouni et al., 2016). We therefore investigated the spectral content of the noise and speech for both AC and BC stimulation. Both the Sennheiser HD650 AC headphones and the B81 BC transducer were measured by a Brüel and Kjær PULSE analyzer equipped with a Head and Torso Simulator (Type 4128C) and an artificial mastoid (Type 4930). The output signal spectrum for both AC and BC stimulation are presented in Figure 5 for the frequency range 160 to 8kHz. In this setup, the target speech and masker are at 0° and the BC stimulation is converted to equal sound pressure. Figure 5 illustrate the spectral difference between AC and BC stimulations in the present study. The AC stimulation (black line) is relatively flat throughout the whole frequency range. In contrast, the BC stimulation is mainly provided at frequencies between 250 and 3150 Hz.

The SRM has two mechanisms that are predominant in different frequency regions: (1) ILDs (e.g., head shadow), which contribute at higher frequencies and (2) ITDs, which contribute at lower frequencies. This means that the high-frequency limitation of BC sound impedes both the general intelligibility due to the reduction of high-frequency sound, but also the benefit of spatial separation due to less use of ILDs.

In situations where the sound signals are equivalent at both ears, such as co-location from the front, BC cross-talk does not disrupt speech understanding. However, when speech and noise are spatially separated, BC cross-talk interferes with the integrity of binaural information, resulting in poorer speech intelligibility compared to AC stimulation, which does not involve such cross-talk. A similar outcome was reported in a previous study (Stenfelt & Zeitouni, 2013).

In contrast to the setup employed by Zeitouni et al. (2016), where the speech originated from the front, this study featured a different configuration wherein the speech source was positioned laterally. In this setup, the current study did not observe any difference in spatial benefit between AC and BC stimulation at a spatial separation of 45°. This

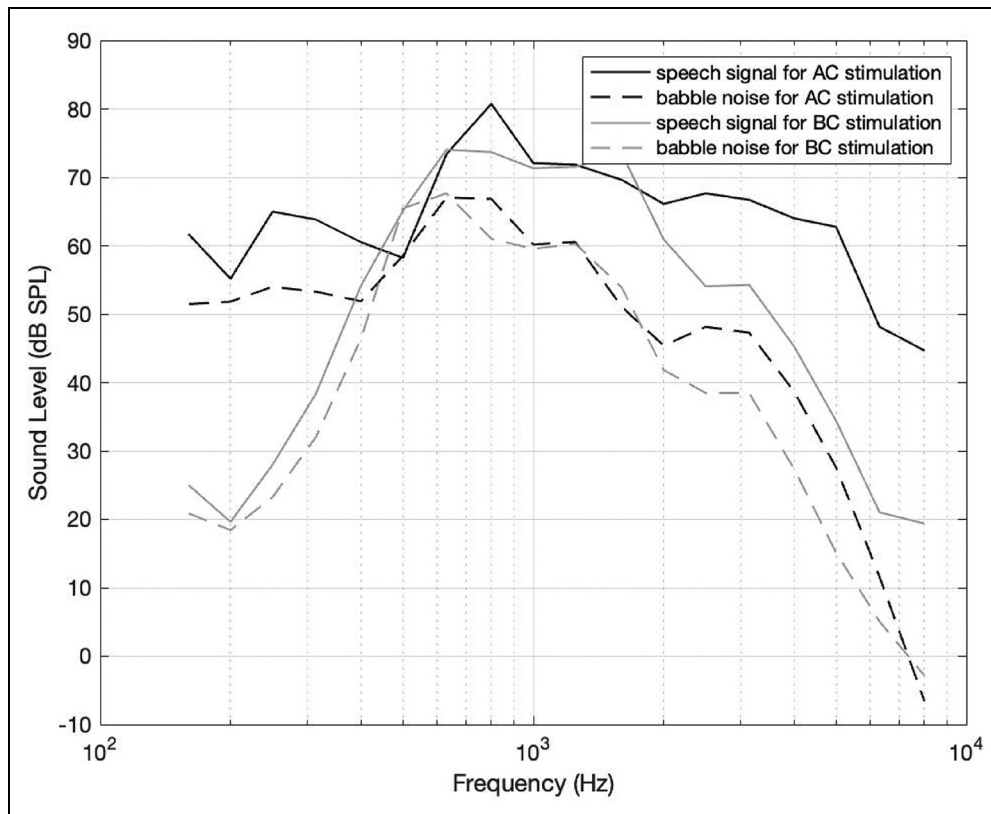


Figure 5. The spectrum of target speech and masker for both the AC and BC stimulation at S0N0 in the frequency range 160 to 8000 Hz. The black solid and dashed lines indicate the target speech and masker for AC stimulation. The gray solid and dashed lines indicate the target speech and masker for BC stimulation. AC = air conduction; BC = bone conduction.

outcome was somewhat unexpected, considering that previous investigations had consistently indicated superior SRM with AC compared to BC stimulation (Stenfelt & Zeitooni, 2013; Zeitooni et al., 2016). However, in the current configuration, the speech was presented near one ear, and the contribution from the opposite ear was minimal due to the acoustic head-shadow. Furthermore, within the initial 45° of spatial separation, alterations in ITD were relatively minor. This observation suggests that the ear on the same side as the speech source played a dominant role in auditory perception. Moreover, binaural processes relying on differences in ITD may not be as effective when the position changes are subtle to the ear with speech stimulation.

As depicted in Figure 2, the SRTs decreased drastically for both AC and BC sound when spatial separations approached 90°. At this position, the masker transitioned from being on the same side as the speech to the opposite side in relation to the speech source. This transition had a profound impact on the TMR at the ipsilateral ear due to the head shadow effect (Edmonds & Culling, 2006; Hawley et al., 2004). As stated above, this effect resulted in an enhanced “better-ear” advantage, where the TMR at the ipsilateral ear was significantly improved. Furthermore, it is worth noting that around 90° of spatial separation, the ITD changes for

the noise source was most pronounced. Consequently, the substantial improvement in SRTs can likely be attributed to the combined influence of the enhanced “better-ear” effect and the masking release resulting from binaural processing.

Figure 3 shows a notable increase in SRM as the spatial separation between target and masker grows, consistent with findings in other investigations (Ahrens et al., 2021; Jelfs et al., 2011; Srinivasan et al., 2016). This phenomenon is a consequence of the augmentation of binaural cues and the better-ear effect, both of which become more pronounced with the increased angular separation between target and masker. Here, BC sound exhibited slightly higher SRM compared to AC sound at spatial separations below 75°. This observation can be attributed to a combination of two mechanisms. Initially, at the S0N0 condition, where sound and masker are collocated, there are limited binaural cues that could lead to a masking release. However, for AC stimulation, owing to the proximity of the stimulation source to the ipsilateral ear, the perception is predominately influenced by the ipsilateral response. In contrast, BC stimulation involves cross-head transmission, which superimposes with HRTF and introduces filtering effects of both ipsilateral and contralateral sound. This superposition can either improve or impair speech intelligibility, depending on the

specific HRTF and cross-head transmission characteristics (Deas et al., 2010; Mattingly et al., 2020; Rowan & Gray, 2008; Stenfelt & Zeitouni, 2013).

In the current study, the BC stimulated SRT is worse than the AC stimulated SRT at SONO, suggesting that the superposition of BC transmission pathways may have detrimental impact on speech intelligibility. The spectrum of speech and masker for both the AC and BC stimulation at SONO were presented in Figure 5. It can be seen from the relatively flat AC stimulation (black lines) throughout the whole frequency range, the important frequency range of the target speech was at 160 Hz to 4 kHz. The BC stimulation is lower than that of AC stimulation, especially at frequencies below 250 Hz. However, as the separation between the speech and noise increased, alterations in ILDs and ITDs for speech and noise modified the filtering effects associated with cross-head transmission. This led to a more substantial enhancement in speech perception for BC than for AC stimulation at smaller separations, as shown in the SRM data presented in Figure 3.

Based on the statistical analysis conducted, no significant difference in SRTs was observed between AC and BC stimulation for spatial separations from 5° to 60°. Even though AC stimulation yielded lower SRTs compared to BC stimulation, these differences only reached statistical significance for spatial separations of 75° and beyond within the current sample. At these separations below 75°, the head-shadow advantage was not yet significant. Instead, the primary factor contributing to the SRTs was the ITD (Culling et al., 2004; Peng & Litovsky, 2021).

As the spatial separation between sound sources increased, the importance of binaural cues became more pronounced. Once again, when a stereo signal is presented bilaterally by BC transducers, cross-head interference may partially or entirely distort the ITD and ILD cues, potentially compromising or erasing the spatial perception (MacDonald et al., 2006). In this study, the specific perceived location of the speech and noise sources were not investigated, leaving uncertainty regarding the participants perceptual experiences. However, previous research exploring sound source localization with bilateral BC stimulation has reported a compressed range of perceived locations skewed toward a frontal position (Denanto et al., 2022; Priwin et al., 2004; Snapp et al., 2020). Moreover, findings by Surendran and Stenfelt (2023) indicated that the time delay associated with cross-head hearing via BC stimulation at the mastoid was on the order of 0.1 ms at 250 Hz. Consequently, due to the distortion of ITD with BC stimulation, the effective separation experienced with BC is less than that with AC stimulation. As a result, the impact of speech and noise separation on the SRTs is expected to be less pronounced for BC compared to AC sound. Indeed, this is reflected in the SRM data in Figure 3, which demonstrates significantly lower SRM values for BC stimulation than those for AC stimulation at separations exceeding 75°. Furthermore, the difference in speech recognition

performance between AC and BC stimulation increased with increasing target-masker angular separation.

The findings in the present study align closely with earlier studies on speech intelligibility. Numerous investigations have explored speech intelligibility, often involving a target talker situated in the front and a single masker positioned at various angles in the horizontal plane (Bronkhorst & Plomp, 1988; Hawley et al., 2004; Peissig & Kollmeier, 1997). These studies have consistently observed comparable trends (Bronkhorst, 2000), typically reporting peak SRM of 10 to 12 dB when the masker position falls between 90° to 120°. In this study, both AC and BC stimulations produced their maximum SRM at approximately 150° spatial separation. In this configuration, the speech source was in close proximity to one ear, while the masker was situated near the opposite ear of the participants. This setup resulted in the highest SRMs, with AC stimulation achieving an SRM of 19.3 dB and BC stimulation achieving an SRM of 13.7 dB. The precise improvement in SRM may vary depending on the specific dataset, but the greater maximum SRMs observed in this study compared to others may be attributed to the nearly maximal ITDs and ILDs generated when speech and noise sources are positioned on opposite sides. Larger ITDs and ILDs are known to enhanced speech intelligibility (Bronkhorst & Plomp, 1988).

In this study, both AC and BC stimulation led to a slight increase in SRTs when the angular separation between target and masker exceeded 150°, resulting in deteriorated speech recognition performance. This could potentially be attributed to the use of nonindividual HRTFs for synthesizing virtual sounds in this study. It is well-established that HRTFs are highly individual-specific, and utilizing nonindividualized HRTFs can significantly impact perception, leading to decreased speech intelligibility, as demonstrated by Ahrens et al. (2021). However, the results regarding SRM in this study differed from the findings in the literature, particularly the work of Plomp et al. (Bronkhorst, 2000, 2015). These differences may be attributed to differences in measurement setups between the two studies, encompassing variations in spatial configuration and the type of masking used. Additionally, the current study deviated from the literature in terms of the spatial separation angle at which SRM was first observed. Previous research demonstrated that SRM could be achieved with very small spatial separations ranging from 2° to 4° for individuals with normal hearing. This was demonstrated by Srinivasan et al. (2016) using a spatial configuration with the target positioned directly in front and symmetrically placed maskers. The variance in acoustic spatial configurations by the target and the masker in our present study could potentially explain the discrepancies observed when compared to the existing literature.

The majority of research on SRM have predominately focused on scenarios where the target and masker were either placed directly in front at 0°, distributed asymmetrically within one hemifield (e.g., maskers at 90° to one side), or

symmetrically divided (e.g., one masker at -90° and the other at 90°) (Ahrens et al., 2021; Hawley et al., 2004; Hess et al., 2018; Stenfelt & Zeitooni, 2013). However, the use of a 90° separation in these previous studies may have resulted in an underestimation of the maximum SRM and limited the generalizability of findings to hearing scenarios where the separation between the target and the masker exceeds 90° . This study expanded the spatial separation to 180° in order to enhance our comprehension of the release from masking phenomenon. This expansion also enabled investigation of the relationship between SRM and angular separation when the target and masker are positioned on opposite sides.

The speech recognition patterns in relation to spatial separation showed consistent trends regardless of the stimulation modality employed. This underscores the potential for enhancing speech comprehension in virtual reality environments by positioning bilateral BC transducers at the mastoid. The similarity between AC and BC stimulation modalities was evident through the following observations:

1. A general decrease in SRT with increasing angular separation between the target and the masker, resulting in a corresponding increase in SRM with greater angular separation.
2. Significant variations in both SRT and SRM when transitioning from masker placement on the ipsilateral side to the contralateral side in relation to the target.
3. The peak performance for speech recognition in both stimulation modalities occurred at the same angular separation.

These findings hold potential implications for spatial sound reproduction using BC devices. For instance, they suggest the possibility of delivering a more predictable

acoustic perception and improved speech intelligibility to users of BC devices in virtual spatial sound reproduction scenarios.

Influence From the Head Shadow

To analyze the effect of improved TMR at the better ear due to the head shadow of the masker, calculation of the speech intelligibility index (SII) (Pavlovic, 2018) was employed. This was accomplished by the method presented in Peng & Litovsky (2021), where the SRT is predicted in the monaural condition with the speech and noise co-located and with the speech and noise spatially separated. The SRT for the SII equal to 0.5 was used as the outcome measure. Moreover, the transducers used will also influence the SRT and to include this in the SII computation, the frequency response of the AC and BC transducers were incorporated in the calculations (Figure 5). The one-third-octave frequency responses of the AC and BC transducers were multiplied with the spectra of speech and noise and used as signal inputs for the computation. According to this, the SRT based on the better ear effect is computed and presented as the SRT for co-located speech and noise minus the SRT for spatially separated speech and noise (Table 2).

A general (no correction for specific dataset) SII of 0.5 was computed for the conditions S0N0, S0N30, S0N60, S0N90, S0N120, S0N150, and S0N180 with both stimulation modalities (Table 2). Since the MHINT has no correction, the SRT predictions by the SII are not expected to equal the experimental data in Figure 2. The results in Table 2 suggest a positive TMR of 6 and 4 dB for AC and BC stimulation for the S0N0 condition. These TMRs change to approximately -13 and -10 dB for AC and BC stimulation at the S0N150 condition, and 3 to 4 dB worse at the S0N180 scenario. The difference between the spatially separated conditions and the collocated condition, that is, the

Table 2. The SRTs of S0N0, S0N30, S0N60, S0N90, S0N120, S0N150, and S0N180 Calculated by the Intelligibility Model.

Stimulation modality	Spatial configurations						
	S0N0	S0N30	S0N60	S0N90	S0N120	S0N150	S0N180
AC	6.6 dB	6.2 dB	5.0 dB	6.1 dB	-7.0 dB	-13.2 dB	-9.2 dB
BC	4.5 dB	5.2 dB	4.8 dB	4.1 dB	-4.6 dB	-9.8 dB	-6.8 dB

AC = air conduction; BC = bone conduction; SRT = speech recognition threshold.

Table 3. The Magnitude of Monaural Head Shadow of S0N0, S0N30, S0N60, S0N90, S0N120, S0N150, and S0N180.

Stimulation modality	Spatial configurations					
	S0N30	S0N60	S0N90	S0N120	S0N150	S0N180
AC	0.4 dB	1.6 dB	0.5 dB	13.6 dB	19.8 dB	15.8 dB
BC			0.4 dB	9.1 dB	14.3 dB	11.3 dB

AC = air conduction; BC = bone conduction.

predicted SRM based on better ear effect alone, is given in Table 3. Consequently, the better ear effect at SON150 and SON180 are approximately 20 and 16 dB for AC stimulation and 14 and 11 dB for BC stimulation (Table 3).

These better ear effects in Table 3 are compared to the SRM data in Figure 3. At separations equal to or $<90^\circ$, the AC SRM increases from around 1 to 7.5 dB. The same SRM increase for BC SRM is 2 to 5 dB. These SRMs are compared to the predicted SRM in Table 3, indicating a significant effect of factors other than the better ear effect alone. However, the better ear effect with the noise at the contralateral side (Table 3) is similar to the tested SRMs in Figure 3. Consequently, with speech and noise on the same side, the better ear effect seems not to contribute to the SRM while with the speech and noise on different sides of the head, the better ear effect seems to dominate the SRM in the current setup.

Conclusions

This article presents experiments that evaluated speech intelligibility under different spatial conditions of bilateral AC stimulation and BC stimulation at the mastoid. The results showed that the angular separation between the target and the masker significantly influenced speech recognition performance. In general, performance improved as the angular separation increased. The most significant improvement in SRM was observed when the target and masker were located on opposite sides of the head.

For BC stimuli, a spatial separation of 10° was the smallest at which a statistically significant SRM was observed. The best speech recognition performance was achieved at a spatial separation of 150° . Similarly, for AC stimulation, the maximum SRM was also achieved at a 150° separation.

The study found that speech recognition was better with AC than with BC stimulation. The difference between AC and BC corresponded to a masking release equivalent to a 2.8° spatial separation with BC stimulation (SRM of 1.9 dB). The maximum SRM obtained by BC stimulation at spatial separations between 150° and 165° was equivalent to the speech recognition performance by AC stimulation at spatial separations between 90° and 105° . Interestingly, when the target and masker were located on the same side, the “better ear effect” did not significantly contribute to the SRM. However, when they were positioned on different sides of the head, this effect dominated the SRM. This observation was consistent across both stimulation modalities.

Acknowledgments

We would like to thank all the participants in this study.


Declaration of Conflicting Interests


The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Science Fund of China (grant number 11974086, 12074403). It was also supported by the Alliance of Guangzhou International Sister-City Universities Research Collaboration Fund of 2023 and Overseas Experts Initiative of Department of Technology and Science of Guangdong Province.

ORCID iDs

Stefan Stenfelt  <https://orcid.org/0000-0003-3350-8997>

Huali Zhou  <https://orcid.org/0000-0002-5710-5271>

Jinxiu Sang  <https://orcid.org/0000-0002-4368-8787>

References

- Ahrens, A., Cuevas Rodriguez, M., & Brimijoin, W. O. (2021). Speech intelligibility with various head-related transfer functions: A computational modelling approach. *JASA Express Letters*, *1*(3), 034401. <https://doi.org/10.1121/10.0003618>
- Arbogast, T. L., Mason, C. R., & Kidd, G. Jr. (2002). The effect of spatial separation on informational and energetic masking of speech. *Journal of the Acoustical Society of America*, *112*(5 Pt 1), 2086–2098. <https://doi.org/10.1121/1.1510141>
- Asp, F., & Reinfeldt, S. (2020). Effects of simulated and profound unilateral sensorineural hearing loss on recognition of speech in competing speech. *Ovid Technologies (Wolters Kluwer Health)*, *41*(2), 411–419. <https://doi.org/10.1097/aud.0000000000000764>
- Best, V., Gallun, F. J., Ihlefeld, A., & Shinn-Cunningham, B. G. (2006). The influence of spatial separation on divided listening. *Journal of the Acoustical Society of America*, *120*(3), 1506–1516. <https://doi.org/10.1121/1.2234849>
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica United with Acustica*, *86*(1), 117–128. <https://doi.org/10.1134/1.29854>
- Bronkhorst, A. W. (2015). The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Attention, Perception & Psychophysics*, *77*(5), 1465–1487. <https://doi.org/10.3758/s13414-015-0882-9>
- Bronkhorst, A. W., & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *Journal of the Acoustical Society of America*, *83*(4), 1508–1516. <https://doi.org/10.1121/1.395906>
- Canale, A., Ndrej, D., Sapino, S., Bianchi, C., Bordino, V., & Albera, A. (2022). Speech in noise with bilateral active bone conduction implant for conductive and mixed hearing loss. *Otology & Neurotology*, *43*(9), 1000–1004. <https://doi.org/10.1097/mao.00000000000003671>
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*(5), 975–979. <https://doi.org/10.1121/1.1907229>
- Colquitt, J. L., Loveman, E., Baguley, D. M., Mitchell, T. E., Sheehan, P. Z., Harris, P., & Welch, K. (2011). Bone-anchored hearing aids for people with bilateral hearing impairment: A systematic review. *Clinical Otolaryngology*, *36*(5), 419–441. <https://doi.org/10.1111/j.1749-4486.2011.02376.x>
- Culling, J. F., Hawley, M. L., & Litovsky, R. Y. (2004). The role of head-induced interaural time and level differences in the speech

- reception threshold for multiple interfering sound sources. *Journal of the Acoustical Society of America*, 116(2), 1057–1065. <https://doi.org/10.1121/1.1772396>
- Deas, R. W., Adamson, R. B., Curran, L. L., Makki, F. M., Bance, M., & Brown, J. A. (2010). Audiometric thresholds measured with single and dual BAHAs transducers: The effect of phase inversion. *International Journal of Audiology*, 49(12), 933–939. <https://doi.org/10.3109/14992027.2010.515621>
- Denanto, F. M., Wales, J., Tideholm, B., & Asp, F. (2022). Differing bilateral benefits for spatial release from masking and sound localization accuracy using bone conduction devices. *Ear and Hearing*, 43(6), 1708–1720. <https://doi.org/10.1097/aud.0000000000001234>
- Dun, C. A., Agterberg, M. J., Cremers, C. W., Hol, M. K., & Snik, A. F. (2013). Bilateral bone conduction devices: Improved hearing ability in children with bilateral conductive hearing loss. *Ear and Hearing*, 34(6), 806–808. <https://doi.org/10.1097/AUD.0b013e318291784e>
- Edmonds, B. A., & Culling, J. F. (2006). The spatial unmasking of speech: Evidence for better-ear listening. *Journal of the Acoustical Society of America*, 120(3), 1539–1545. <https://doi.org/10.1121/1.2228573>
- Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *Journal of the Acoustical Society of America*, 106(6), 3578–3588. <https://doi.org/10.1121/1.428211>
- Glyde, H., Buchholz, J. M., Dillon, H., Cameron, S., & Hickson, L. (2013). The importance of interaural time differences and level differences in spatial release from masking. *Journal of the Acoustical Society of America*, 134(2), EL147–EL152. <https://doi.org/10.1121/1.4812441>
- Hawley, M. L., Litovsky, R. Y., & Colburn, H. S. (1999). Speech intelligibility and localization in a multi-source environment. *Journal of the Acoustical Society of America*, 105(6), 3436–3448. <https://doi.org/10.1121/1.424670>
- Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *Journal of the Acoustical Society of America*, 115(2), 833–843. <https://doi.org/10.1121/1.1639908>
- Heath, E., Dawoud, M. M., Stavrakas, M., & Ray, J. (2022). The outcomes of bilateral bone conduction hearing devices (BCHD) implantation in the treatment of hearing loss: A systematic review. *Cochlear Implants International*, 23(2), 95–108. <https://doi.org/10.1080/14670100.2021.2001926>
- Hess, C. L., Misurelli, S. M., & Litovsky, R. Y. (2018). Spatial release from masking in 2-year-olds with normal hearing and with bilateral cochlear implants. *Trends in Hearing*, 22, 2331216518775567. <https://doi.org/10.1177/2331216518775567>
- Jelfs, S., Culling, J. F., & Lavandier, M. (2011). Revision and validation of a binaural model for speech intelligibility in noise. *Hearing Research*, 275(1-2), 96–104. <https://doi.org/10.1016/j.heares.2010.12.005>
- Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61(5), 1337–1351. <https://doi.org/10.1121/1.381436>
- Kang, H. (2021). Sample size determination and power analysis using the G*Power software. *Journal of Educational Evaluation for Health Professions*, 18, 17. <https://doi.org/10.3352/jehp.2021.18.17>
- Kidd, G. Jr., Mason, C. R., Best, V., & Marrone, N. (2010). Stimulus factors influencing spatial release from speech-on-speech masking. *Journal of the Acoustical Society of America*, 128(4), 1965–1978. <https://doi.org/10.1121/1.3478781>
- Kidd, G. Jr., Mason, C. R., Rohtla, T. L., & Deliwala, P. S. (1998). Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *Journal of the Acoustical Society of America*, 104(1), 422–431. <https://doi.org/10.1121/1.423246>
- Levitt, H. C. C. H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(2B), 467–477. <https://doi.org/10.1121/1.1912375>
- Lindeman, R. W., Noma, H., & de Barros, P. G. (2008, March). An empirical study of hear-through augmented reality: Using bone conduction to deliver spatialized audio [Paper presentation]. 2008 IEEE Virtual Reality Conference (pp. 35–42).
- Litovsky, R. Y. (2012). Spatial release from masking. *Acoustics Today*, 8(2), 18–25. <https://doi.org/10.1121/1.4729575>
- MacDonald, J. A., Henry, P. P., & Letowski, T. R. (2006). Spatial audio through a bone conduction interface. *International Journal of Audiology*, 45(10), 595–599. <https://doi.org/10.1080/14992020600876519>
- Maier, L. T., Agha-Mir-Salim, P., Agterberg, M. J. H., Anagiotos, A., Arndt, S., Briggs, R. B. S., Caversaccio, M., Dahm, M., Dalhoff, E., Devèze, A., Ebrahimi-Madiseh, A., Frayse, B., Frenzel, H., Gavilán, J., Ghoncheh, M., Hakansson, B. E. V., Hodgetts, W., Hol, M., Holland, J., Holmberg, M., Huber, A. M., Jenkins, H., Katiri, R., Kheirkhah, K., Koitschev, A., Kompis, M., Lanting, C., Lassaletta, L., Lerut, B., Leuwer, R., Linder, T., Löwenheim, H., Lustig, L., Mandavia, R., Manrique, M., Martins, J. H., Mertens, G., Mlynski, R., Mojallal, H., Monini, S., Monksfield, P., Müller, A., Mylanus, E., Nakajima, H., Neudert, M., Offeciers, E., Piffner, F., Pietsch, M., Plontke, S. K., Prenzler, N., Profant, M., Rahne, T., Rajan, G., Ratuszniak, A., Raufer, S., Ray, J., Reinfeldt, S., Röösl, C., Rosenbom, T., Salcher, R., Schönermark, M., Schwab, B., Skarżyński, H., Skarżyński, P. H., Snapp, H., Sprinzel, G., Spearman, M., Stenfelt, S., Stieger, C., Tringali, S., Truy, E., Tysome, J., Van de Heyning, P., Verhaert, N., Wesarg, T., Westerkull, P., Wollenberg, B., Zahnert, T., Zarowski, A., & Snik, A. (2022). Consensus statement on bone conduction devices and active middle ear implants in conductive and mixed hearing loss. *Otology & Neurotology*, 43(5), 513–529. <https://doi.org/10.1097/mao.0000000000003491>
- Marrone, N., Mason, C. R., & Kidd, G. Jr. (2008a). The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *Journal of the Acoustical Society of America*, 124(5), 3064–3075. <https://doi.org/10.1121/1.2980441>
- Marrone, N., Mason, C. R., & Kidd, G. (2008). Tuning in the spatial dimension: Evidence from a masked speech identification task. *Journal of the Acoustical Society of America*, 124(2), 1146–1158. <https://doi.org/10.1121/1.2945710>
- Mattingly, J. K., Banakis Hartl, R. M., Jenkins, H. A., Tollin, D. J., Cass, S. P., & Greene, N. T. (2020). A comparison of intracochlear pressures during ipsilateral and contralateral stimulation with a bone conduction implant. *Ear and Hearing*, 41(2), 312–322. <https://doi.org/10.1097/aud.0000000000000758>
- McBride, M., Tran, P., Pollard, K. A., Letowski, T., & McMillan, G. P. (2015). Effects of bone vibrator position on auditory spatial

- perception tasks. *Human Factors*, 57(8), 1443–1458. <https://doi.org/10.1177/0018720815596272>
- McLeod, R. W. J., & Culling, J. F. (2020). Unilateral crosstalk cancellation in normal hearing participants using bilateral bone transducers. *Journal of the Acoustical Society of America*, 148(1), 63–72. <https://doi.org/10.1121/10.0001529>
- Pavlovic, C. (2018). SII—speech intelligibility index standard: ANSI S3. 5 1997. *Journal of the Acoustical Society of America*, 143(3_Supplement), 1906–1906. <https://doi.org/10.1121/1.5036206>
- Peissig, J., & Kollmeier, B. (1997). Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners. *Journal of the Acoustical Society of America*, 101(3), 1660–1670. <https://doi.org/10.1121/1.418150>
- Peng, Z. E., & Litovsky, R. Y. (2021). Novel approaches to measure spatial release from masking in children with bilateral cochlear implants. *Ear and Hearing*, 43(1), 101–114. <https://doi.org/10.1097/aud.0000000000001080>
- Priwin, C., Stenfelt, S., Granström, G., Tjellström, A., & Håkansson, B. (2004). Bilateral bone-anchored hearing aids (BAHAs): An audiometric evaluation. *The Laryngoscope*, 114(1), 77–84. <https://doi.org/10.1097/00005537-200401000-00013>
- Qin, X., & Usagawa, T. (2017). Frequency characteristics of bone conduction actuators—measurements of loudness and acceleration. *Applied Acoustics*, 126, 19–25. <https://doi.org/10.1016/j.apacoust.2017.05.007>
- Reinfeldt, S., Ostli, P., Håkansson, B., & Stenfelt, S. (2010). Hearing one's own voice during phoneme vocalization—transmission by air and bone conduction. *Journal of the Acoustical Society of America*, 128(2), 751–762. <https://doi.org/10.1121/1.3458855>
- Reinfeldt, S., Stenfelt, S., Good, T., & Håkansson, B. (2007). Examination of bone-conducted transmission from sound field excitation measured by thresholds, ear-canal sound pressure, and skull vibrations. *Journal of the Acoustical Society of America*, 121(3), 1576–1587. <https://doi.org/10.1121/1.2434762>
- Reinfeldt, S., Stenfelt, S., & Håkansson, B. J. H. R. (2013). Estimation of bone conduction skull transmission by hearing thresholds and ear-canal sound pressure. *Hearing Research*, 299, 19–28. <https://doi.org/10.1016/j.heares.2013.01.023>
- Rowan, D., & Gray, M. (2008). Lateralization of high-frequency pure tones with interaural phase difference and bone conduction. *International Journal of Audiology*, 47(7), 404–411. <https://doi.org/10.1080/14992020802006055>
- Snapp, H., Vogt, K., & Agterberg, M. J. H. (2020). Bilateral bone conduction stimulation provides reliable binaural cues for localization. *Hearing Research*, 388, 107881. <https://doi.org/10.1016/j.heares.2019.107881>
- Srinivasan, N. K., Jakien, K. M., & Gallun, F. J. (2016). Release from masking for small spatial separations: Effects of age and hearing loss. *Journal of the Acoustical Society of America*, 140(1), E173–E178. <https://doi.org/10.1121/1.4954386>
- Stenfelt, S. (2005). Bilateral fitting of BAHAs and BAHA fitted in unilateral deaf persons: Acoustical aspects. *International Journal of Audiology*, 44(3), 178–189. <https://doi.org/10.1080/14992020500031561>
- Stenfelt, S. (2007). Simultaneous cancellation of air and bone conduction tones at two frequencies: Extension of the famous experiment by von Békésy. *Hearing Research*, 225(1-2), 105–116. <https://doi.org/10.1016/j.heares.2006.12.009>
- Stenfelt, S. (2011). Acoustic and physiologic aspects of bone conduction hearing. *Advances in Otorhino-Laryngology*, 71, 10–21. <https://doi.org/10.1159/000323574>
- Stenfelt, S. (2012). Transcranial attenuation of bone-conducted sound when stimulation is at the mastoid and at the bone conduction hearing aid position. *Otology & Neurotology*, 33(2), 105–114. <https://doi.org/10.1097/MAO.0b013e31823e28ab>
- Stenfelt, S. (2020). Investigation of mechanisms in bone conduction hyperacusis with third window pathologies based on model predictions. *Frontiers in Neurology*, 11, 966. <https://doi.org/10.3389/fneur.2020.00966>
- Stenfelt, S., & Goode, R. L. (2005). Bone-conducted sound: Physiological and clinical aspects. *Otology & Neurotology*, 26(6), 1245–1261. <https://doi.org/10.1097/01.mao.00000187236.10842.d5>
- Stenfelt, S., & Prodanovic, S. (2022). Simulation of soft tissue stimulation—indication of a skull bone vibration mechanism in bone conduction hearing. *Hearing Research*, 418, 108471. <https://doi.org/10.1016/j.heares.2022.108471>
- Stenfelt, S., Puria, S., Hato, N., & Goode, R. L. (2003). Basilar membrane and osseous spiral lamina motion in human cadavers with air and bone conduction stimuli. *Hearing Research*, 181(1-2), 131–143. [https://doi.org/10.1016/s0378-5955\(03\)00183-7](https://doi.org/10.1016/s0378-5955(03)00183-7)
- Stenfelt, S., & Zeitooni, M. (2013). Binaural hearing ability with mastoid applied bilateral bone conduction stimulation in normal hearing subjects. *Journal of the Acoustical Society of America*, 134(1), 481–493. <https://doi.org/10.1121/1.4807637>
- Surendran, S., Prodanovic, S., & Stenfelt, S. (2023). Hearing through bone conduction headsets. *Trends in Hearing*, 27, 23312165231168741. <https://doi.org/10.1177/23312165231168741>
- Surendran, S., & Stenfelt, S. (2023). Inter-aural separation during hearing by bilateral bone conduction stimulation. *Hearing Research*, 437, 108852. <https://doi.org/10.1016/j.heares.2023.108852>
- Voong, T. M., & Oehler, M. (2019, March). Auditory spatial perception using bone conduction headphones along with fitted head related transfer functions [Paper presentation]. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (pp. 1211–1212). IEEE.
- Wang, J., Chen, Y., Stenfelt, S., Sang, J., Li, X., & Zheng, C. (2023). Analysis of cross-talk cancellation of bilateral bone conduction stimulation. *Hearing Research*, 434, 108781. <https://doi.org/10.1016/j.heares.2023.108781>
- Wong, L. L., Soli, S. D., Liu, S., Han, N., & Huang, M. W. (2007). Development of the mandarin hearing in noise test (MHINT). *Ear and Hearing*, 28(2 Suppl), 70s–74s. <https://doi.org/10.1097/AUD.0b013e31803154d0>
- Zeitooni, M., Mäki-Torkko, E., & Stenfelt, S. (2016). Binaural hearing ability with bilateral bone conduction stimulation in subjects with normal hearing: Implications for bone conduction hearing aids. *Ear and Hearing*, 37(6), 690–702. <https://doi.org/10.1097/aud.0000000000000336>
- Zwislocki, J. (2005). Acoustic attenuation between the ears. *Journal of the Acoustical Society of America*, 25(4), 752–759. <https://doi.org/10.1121/1.1907171>