

Effect of Microphone Configuration and Sound Source Location on Speech Recognition for Adult Cochlear Implant Users with Current-Generation Sound Processors

Robert T. Dwyer¹ Jillian Roberts¹ René H. Gifford^{1,2}

¹Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, Tennessee

Address for correspondence Robert T. Dwyer, AuD, Robert.Dwyer@vumc.org.

²Department of Otolaryngology, Vanderbilt University Medical Center, Nashville, Tennessee

J Am Acad Audiol 2020;31:578–589.

Abstract

Background Microphone location has been shown to influence speech recognition with a microphone placed at the entrance to the ear canal yielding higher levels of speech recognition than top-of-the-pinna placement. Although this work is currently influencing cochlear implant programming practices, prior studies were completed with previous-generation microphone and sound processor technology. Consequently, the applicability of prior studies to current clinical practice is unclear.

Purpose To investigate how microphone location (e.g., at the entrance to the ear canal, at the top of the pinna), speech-source location, and configuration (e.g., omnidirectional, directional) influence speech recognition for adult CI recipients with the latest in sound processor technology.

Research Design Single-center prospective study using a within-subjects, repeated-measures design.

Study Sample Eleven experienced adult Advanced Bionics cochlear implant recipients (five bilateral, six bimodal) using a Naída CI Q90 sound processor were recruited for this study.

Data Collection and Analysis Sentences were presented from a single loudspeaker at 65 dBA for source azimuths of 0°, 90°, or 270° with semidiffuse noise originating from the remaining loudspeakers in the R-SPACE array. Individualized signal-to-noise ratios were determined to obtain 50% correct in the unilateral cochlear implant condition with the signal at 0°. Performance was compared across the following microphone sources: T-Mic 2, integrated processor microphone (formerly behind-the-ear mic), processor microphone + T-Mic 2, and two types of beamforming: monaural, adaptive beamforming (UltraZoom) and binaural beamforming (StereoZoom). Repeated-measures analyses were completed for both speech recognition and microphone output for each microphone location and configuration as well as sound source location. A two-way analysis of variance using mic and azimuth as the independent variables and output for pink noise as the dependent variable was used to characterize the acoustic output characteristics of each microphone source.

Keywords

- ▶ cochlear implant
- ▶ beamforming
- ▶ directional microphone
- ▶ microphone location
- ▶ R-SPACE
- ▶ T-Mic
- ▶ SNR
- ▶ speech recognition

received
March 19, 2019
accepted after revision
January 25, 2020
published online
April 27, 2020

Copyright © 2020 by the American Academy of Audiology. All rights reserved. Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA.
Tel: +1(212) 760-0888.

DOI <https://doi.org/10.1055/s-0040-1709449>.
ISSN 1050-0545.

Results No significant differences in speech recognition across omnidirectional mic location at any source azimuth or listening condition were observed. Secondary findings were (1) omnidirectional microphone configurations afforded significantly higher speech recognition for conditions in which speech was directed to $\pm 90^\circ$ (when compared with directional microphone configurations), (2) omnidirectional microphone output was significantly greater when the signal was presented off-axis, and (3) processor microphone output was significantly greater than T-Mic 2 when the sound originated from 0° , which contributed to better aided detection at 2 and 6 kHz with the processor microphone in this group.

Conclusions Unlike previous-generation microphones, we found no statistically significant effect of microphone location on speech recognition in noise from any source azimuth. Directional microphones significantly improved speech recognition in the most difficult listening environments.

Despite advances in cochlear implant (CI) technology, recipients continue to report difficulty recognizing speech under noisy, real-world conditions. CI recipients demonstrate significant decrements in speech recognition in noise as compared with quiet conditions even in generous signal-to-noise ratios (SNRs) such as +10 dB. As an example, mean AzBio³⁴ sentence recognition at +10 dB SNR ranges from 39 to 71% for unilateral CI-alone conditions and from 62 to 81% for bimodal listeners.¹⁻⁴ In a more realistic SNR of +5 dB, mean AzBio sentence recognition ranges from 22 to 57% for the unilateral CI alone, 27 to 49% for bimodal listeners, and 37 to 66% for bilateral CI listeners.^{1,2,5-7} For comparison, adults (age: 21-79) with normal hearing score 95 to 99%, on average, for AzBio sentence recognition at +5 dB^{6,8} and children with normal hearing score 98% for pediatric AzBio sentences at +5 dB SNR.⁹ The results of the aforementioned works, and reports from recipients themselves, illustrate the difficulty CI recipients face while listening in environments with competing noise, even with a favorable SNR not often available in the real world.

Effects of Directional Microphones

While noise reduction (e.g., ClearVoice) and external accessories (e.g., remote microphone systems) have traditionally been recommended to CI recipients to improve speech recognition under less than ideal listening conditions, more recently directional microphone technology has been incorporated into the latest generation of CI sound processors. The basic principle of directional microphones is the same across manufacturers. The signal arrives at two or more spatially separated microphones. Because the signal arrives at each microphone at differing times, signal processing can be used to take advantage of phase differences to shape a variety of sensitivity patterns. These patterns can be used to fix the maximum point of attenuation (i.e., null) or continuously change it to suppress a noise source as the location of the noise changes.

UltraZoom is a monaural, adaptive beamformer used in the Advanced Bionics (AB) Naída CI Q70/90 sound processor. While previous works using a variety of testing protocols (e.g., outcome measures, speech materials, noise types, noise sources,

additional signal processing, etc.) make expected benefits difficult to generalize, UltraZoom has been shown to provide benefit to CI recipients. Dorman et al¹⁰ presented female-voiced target sentences from the pediatric AzBio sentence corpus¹¹ from 0° , while male-voiced distractors were presented from $\pm 90^\circ$. The target signal was presented at 60 dB sound pressure level (SPL). Individual performance in the omnidirectional listening condition was driven down to 30 to 60% by adjusting the level of the male talkers. This SNR was then used for additional testing in noise using UltraZoom. A 31% benefit in speech recognition over omnidirectional performance in 10 adults (fit unilaterally) was observed.

Mosnier et al¹² investigated the effect of UltraZoom on speech recognition in noise in 21 adult recipients. Speech in noise performance was measured using the Matrix sentence test in French.¹³ The group used a noncorrelated speech-shaped noise presented at 65 dBA from three speakers ($\pm 90^\circ$ and 180°). The speech signal was adaptively adjusted to arrive at a presentation level where the participant understood 50% of the target signal (i.e., speech reception threshold, SRT). In this evaluation, UltraZoom yielded a median improvement of 3.6 dB in SRT as compared with the omnidirectional mode. Significant subjective improvement on the background noise and aversiveness subscales of the APHAB (Abbreviated Profile of Hearing Aid Benefit) questionnaire was also observed.

Until Holder et al,¹⁴ previous work showing the benefit of UltraZoom had only been completed in adult populations. The work of Holder et al¹⁴ is the first pediatric study completed with UltraZoom. The group presented AzBio sentence materials from 0° in the R-SPACE proprietary restaurant noise at a +5 dB SNR. UltraZoom provided, on average, a 15-percentage point improvement in speech recognition in noise over the T-Mic 2 in a group of nine pediatric CI recipients.

While investigations of binaural beamforming in CIs are not new, Buechner et al¹⁵ were the first to use a system in which the experimental setup did not require a central processing unit. In this experiment, Phonak Ambra hearing aids (HAs) were connected to the auxiliary input of the AB Harmony processor. The wireless communication between the HAs

allowed for binaural beamforming using the four microphones across the two devices. To measure SRT, Oldenburg sentences were presented in uncorrelated noise from five loudspeakers surrounding the listener with an overall level of 65 dB SPL. The level of the target signal was varied until arriving at SRT. They noted a 7.1-dB improvement in SRT with binaural beamforming over the omnidirectional microphone listening condition.

Binaural beamforming has since been made commercially available as “StereoZoom” and is available to Naída CI Q90 recipients. StereoZoom combines the signal from two independent dual-microphone systems (e.g., bilateral Naída CI Q90 or Naída CI Q90 + Naída Link HA). The monaural beamforming algorithm is first processed in each device independently. The resulting ear-specific directional signals are transmitted to the contralateral ear where the signal is combined with the contralateral, ear-specific, directional signal. The result is a much narrower (and fixed) four-microphone beamformer.

The first evaluations of StereoZoom are just beginning to be reported in the literature. Ernst et al¹⁶ investigated the speech recognition benefit of StereoZoom over T-Mic 2 in 10 bilateral CI and 10 bimodal CI recipients. SRTs were measured using Oldenburg sentences in noise with the speech signal always presented from 0°. Speech-shaped noise was fixed at 65 dB SPL from the five remaining loudspeakers. Two different speaker arrangements were used. Setup A positioned the speakers at 60° intervals around the listener. Setup B arranged the loudspeakers at ±30°, ±60°, and one speaker at 180°. For the bilateral group, a significant advantage of StereoZoom was noted over T-Mic 2 in test setup A (5.2 dB) and in test setup B (3.4 dB). There was also a significant advantage of StereoZoom over UltraZoom in test setup B (1.4 dB). For bimodal participants, a significant advantage of StereoZoom over the T-Mic 2 was observed in both test setup A (4.6 dB) and test setup B (2.6 dB). StereoZoom in bimodal listeners offered a significant advantage over UltraZoom in both test setup A (1.3 dB) and test setup B (1.2 dB). Vroegop et al¹⁷ also showed a 4.7-dB advantage in SRT of StereoZoom over omnidirectional microphone sources in bimodal listeners.

While both Ernst et al¹⁶ and Vroegop et al¹⁷ evaluated bimodal and bilateral AB recipients, their designs are quite different when compared with this current work. Ernst et al¹⁶ employed both a 5- and 6-speaker setup to investigate the advantage of directional microphone technology over only the T-Mic 2 omnidirectional microphone. Investigators used a speech-shaped noise, which is not typical of the noise that CI listeners encounter in their everyday lives. Additionally, because the investigators were interested in the impact of directionality on speech recognition, the speech signal was presented from the front of the listener only. While Vroegop et al¹⁷ did present the signal slightly off-axis in some listening conditions (i.e., 45°), the speech signal the investigators used was presented at 70 dB SPL, which according to Pearsons et al,¹⁸ is well above what the researchers called “raised speech” for male (65 dB SPL) and for female (63 dB SPL) talkers. These design differences motivated the current study which seeks to add to the current body of literature by evaluating directional microphone technology using a semidiffuse noise source and signals that, in addition to arriving from the front of

the listener at a more appropriate conversational level, arrive off-axis as would occur in a more realistic setting.

Effects of Microphone Location

One additional solution to improve speech recognition in noise that has previously been reported in HA listeners, and only more recently explored in CI listeners, is the influence of microphone location on SNR and, consequently, speech recognition in environments with less favorable SNRs. Mantokoudis et al¹⁹ studied the effect of microphone location on speech recognition performance in adult implant recipients for both in the canal (ITC) and a behind-the-ear (BTE) microphone locations. While the differences in speech recognition did not reach statistical significance, the ITC provided a 3.0-dB improvement in SRT.

In another study, Aronoff et al²⁰ obtained SRTs for normal-hearing listeners using head-related transfer functions (HRTFs) obtained from the BTE microphone and the T-Mic in the AB Harmony sound processor. Results indicated that the T-Mic, which sits just outside the entrance to the ear canal, yielded a 2-dB advantage in SRT as compared with the BTE microphone. Additionally, the SRT obtained from the T-Mic HRTF was not significantly different from the KEMAR HRTF, which represents unaided acoustic hearing including pinna effects.

Gifford and Revit²¹ also investigated the effect of microphone location on speech recognition. SRTs were obtained from 14 adult CI recipients in semidiffuse noise using the R-SPACE sound simulation system. They observed a 4.4-dB improvement in SRT with the Harmony T-Mic as compared with the BTE microphone.

Until the work of Kolberg and colleagues,²² no study had specifically investigated the effect of CI microphone location for a signal (e.g., speech) originating from various source azimuths—as is typically encountered in real-world listening environments such as in a small-group gathering or at a dinner party. Kolberg et al²² found that the output of the AB Harmony BTE microphone was 5 dB less from 1,500 to 4,500 Hz for signals presented at 0° as compared with 90° (i.e., toward the processor), but that the Harmony T-Mic output was essentially equivalent for sources originating from 0° and 90°. As shown below, microphone location significantly impacted sentence recognition as a function of source azimuth with the Harmony T-Mic yielding the highest performance for speech from 0° azimuth. This study highlighted the benefit afforded by the T-Mic, which is placed at the entrance of the ear canal as compared with a BTE microphone placement, which rests on the top of the pinna.

In contrast to these findings, Dwyer et al²³ found speech recognition in noise was greater when the speech signal was presented from 90° (33.1%) than when the speech signal was presented from the front of the listener (17.6%) while using the latest-generation sound processor (Naída CI Q90) and T-Mic (i.e., T-Mic 2). They attributed this phenomenon to a partial head-shadow (they termed “face-shadow”) or the difference in speech recognition in unilateral listeners when the signal is presented to the device side of the listener as compared with the front of the listener where the signal could be impeded prior to reaching the microphone on implanted ear. While

these findings contradict previous work, they highlight the importance of continuing to evaluate sound processor technology as it evolves.

Since the publication of the aforementioned studies, CI sound processor technology has advanced and sound processors are now equipped with monaural adaptive and binaural beamforming (when used in conjunction with a second compatible device) in addition to traditional omnidirectional microphones. Thus, the primary aims of this study were to determine the effects of microphone location (e.g., at the entrance to the ear canal, at the top of the pinna), speech-source location, and listening configuration (e.g., omnidirectional, directional) on speech recognition for adult CI recipients with the latest in sound processor technology. On the basis of previous literature examining the effects of microphone location in CIs and in HAs, our hypotheses were (1) the T-Mic 2 would result in significantly higher speech recognition than all other omnidirectional microphone locations, (2) speech recognition would be greatest for omnidirectional microphones with speech originating at 90°, and (3) UltraZoom and StereoZoom would afford significant benefit over omnidirectional configurations for speech recognition in noise with speech at 0°.

Methods

Participants

Eleven experienced (5 bilateral CI, 6 bimodal) adult CI recipients implanted with the AB CI system (Valencia, CA) participated in this study conducted in accordance with local university institutional review board approval (IRB number: 131315). Participants ranged in age from 35 years to 71 years (mean = 55.5 years). All participants were postlingually deafened and were required to have at least 6 months experience with his or her CI(s) to meet inclusion criteria. All bimodal participants were fitted with a Phonak Naída Link HA in the unimplanted ear. All HAs were fitted to NAL-NL2 targets using real-ear measures.²⁴ Audiometric thresholds for the unimplanted ear are shown in ►Fig. 1. All participants were wearing Naída CI Q90 sound processor (s). See ►Table 1 for additional demographic information.

Stimuli

Testing was completed in a single-walled sound booth using the Revitronix (Braintree, VT) R-SPACE sound simulation system. As described in detail in previous studies,^{25,26} the R-SPACE uses eight loudspeakers arranged at 45° intervals with each speaker positioned 24 inches from the listener's head to simulate a realistic restaurant environment.

The Texas Instruments Massachusetts Institute of Technology (TIMIT) sentences²⁷⁻³¹ were randomly presented from a single speaker located at 0°, 90°, or 270°. The R-SPACE proprietary restaurant noise was presented from the remaining seven speakers. Previous work by Dorman et al^{29,30} and Loizou et al²⁸ created a subset of 34 lists (20 sentences per list) equated for equal intelligibility. We used the 29 TIMIT lists demonstrated to have the highest test-retest reliability as demonstrated by King et al.³¹ In the

current study, 10 groups composed of three lists each were created (e.g., 60 sentences per group). The TIMIT sentences, spoken by both male and female speakers representing eight different American English dialects, were presented at 65 dBA for all conditions in the current experiment.

Procedure

Individual SNR was measured using a single TIMIT list to achieve approximately 50% correct in the unilateral implant (or best CI for bilateral CI listeners) listening condition with speech originating from a single loudspeaker at 0°, and the R-SPACE restaurant noise presented from the remaining seven speakers. The final SNR used ranged from +2 to +15 dB with a mean of +7.64 dB (►Table 1). The individually determined SNR was used for all testing going forward. During testing, participants were instructed to face the speaker placed at 0°, regardless of the speech signal azimuth.

Two list groups (120 TIMIT sentences total) were presented for each microphone source. Testing was completed with T-Mic 2, the integrated processor microphone (formerly referred to as BTE mic), the processor microphone + T-Mic 2 (formerly referred to as 50/50), as well as the UltraZoom and StereoZoom. As in Kolberg et al,²² the processor microphone + T-Mic 2 mixing condition was included because, at the time of experimentation, this was the default microphone setting in the AB clinical programming software (SoundWave 3.0). As a result, this microphone source mixing is common in everyday use for many AB implant recipients. Sentences were presented randomly (40 sentences per azimuth) in the best-aided listening configuration (bilateral or bimodal). Scores were recorded as a percent correct for each azimuth (0°, 90°, or 270°). The order of the microphone source and the list groups used for each source was randomized by the test administrators prior to experimentation.

For the physical sound-level measurements, a Naída CI Q90 was fitted on a KEMAR acoustic mannequin. Microphone output was recorded for pink noise presented at 0° and 90° for the T-Mic 2 and processor microphone because these two microphone sources are used most often clinically and these were the microphone sources investigated in a previous study.²² These measures were completed not only to compare the output of each microphone source and azimuth, but also because there are no known published data detailing the output response characteristics of the microphone sources on the Naída CI Q90 processor.

Results

Bimodal CI Speech Recognition

Mean bimodal performance is summarized in ►Table 2. Speech recognition was poorest for all omnidirectional mic configurations when speech originated from the HA side (90° or 270°). Mean performance for speech directed toward the HA ear was 58.6% for the processor microphone, 58.5% for the T-Mic 2, and 62.0% for processor microphone + T-Mic 2. No detriment to bimodal speech recognition was observed when speech was presented to the HA ear with UltraZoom (59.3%) or StereoZoom (59.0%).

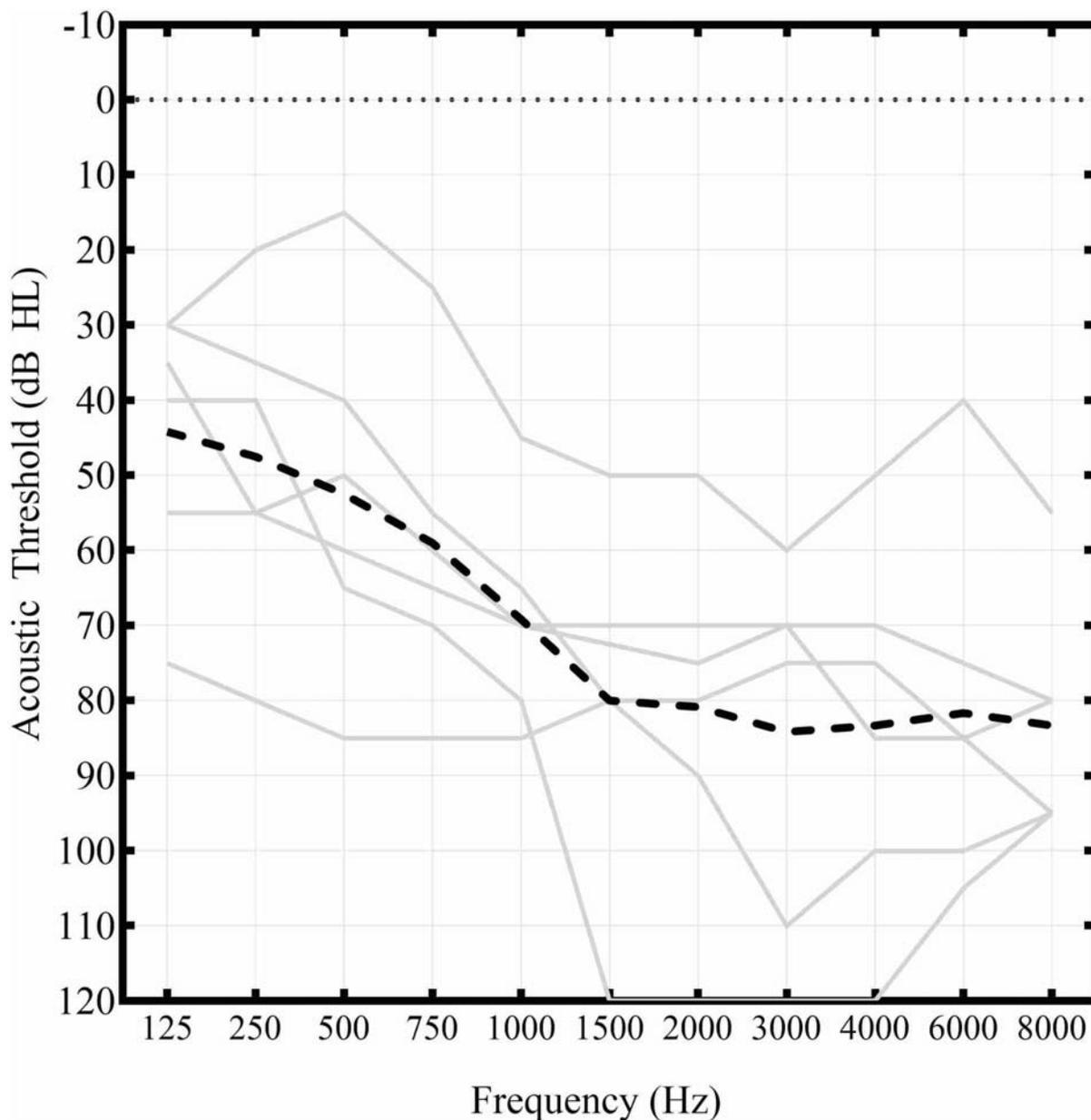


Fig. 1 Unaided acoustic thresholds as a function of frequency for the non-CI ear in the bimodal group. Mean acoustic thresholds are shown by the dashed line. CI, cochlear implant.

Mean bimodal speech recognition for speech originating from the CI side (90° or 270°) was 70.1% for the processor microphone, 65.8% for the T-Mic 2, and 68.8% for processor microphone + T-Mic 2. Beamforming resulted in significantly poorer speech recognition when speech was presented to the CI ear as compared with all omnidirectional mics. Specifically, with speech directed toward the CI, bimodal speech recognition was 38.0 and 31.5% with UltraZoom and StereoZoom, respectively.

Mean bimodal speech recognition for speech originating from 0° was 68.3% for the processor microphone, 65.9% for the T-Mic 2, and 63.4% for the processor microphone + T-Mic 2. Beamforming resulted in higher speech recognition scores with speech at 0° as compared with all omnidirectional mics. Specifically, with speech at 0° , bimodal speech recognition was 77.9 and 78.9% with UltraZoom and StereoZoom, respectively.

Bilateral CI Speech Recognition

Mean bilateral speech recognition is summarized in ► **Table 2**. Speech recognition scores for speech directed toward the *poorer CI ear* (90° or 270°) were 49.7% for the processor microphone, 47.0% for the T-mic 2, and 46.6% for processor microphone + T-Mic 2. Beamforming resulted in poorer speech recognition as compared with all omnidirectional mic conditions with speech at 90° or 270° . Specifically, mean bilateral CI speech recognition scores for speech directed to the poorer CI ear were 22.4 and 24.7% for UltraZoom and StereoZoom, respectively.

Mean bilateral CI speech recognition scores for speech directed toward the *better CI ear* (90° or 270°) were 59.6% for the processor microphone, 61.9% for the T-Mic 2, and 63.7% for processor microphone + T-Mic 2. Beamforming resulted in poorer speech recognition as compared with all omnidirectional mic conditions with speech at 90° or 270° . Specifically,

Table 1 Demographic information including age at testing, gender, recipient type, signal-to-noise ratio (SNR) used for testing, and years of CI experience

Participant	Age	Gender	Recipient type	SNR used for testing	Years of CI experience	
					First CI	Second CI
S1	35	Male	Bilateral	2	13.25	13.17
S2	69	Male	Bilateral	15	4.08	2.50
S3	50	Male	Bilateral	11	6.17	5.08
S4	50	Female	Bilateral	10	0.83	0.58
S5	49	Male	Bilateral	3	1.58	1.08
S6	64	Male	Bimodal	9	3.75	
S7	68	Male	Bimodal	13	1.17	
S8	45	Male	Bimodal	10	0.58	
S9	34	Female	Bimodal	4	2.92	
S10	69	Male	Bimodal	2	1.08	
S11	71	Male	Bimodal	5	9.25	
Mean	55.50	2 Female	–	7.64	4.06	4.48

Abbreviation: CI, cochlear implant.

Table 2 Average percentage correct for bimodal and bilateral listeners for each microphone source and speech signal location

CI microphone source	Bimodal listeners		
	Signal location		
	HA	Front	CI
Processor Mic	58.6%	68.3%	70.1%
T-Mic 2	58.5%	65.9%	65.8%
P-Mic + T-Mic 2	62.0%	63.4%	68.8%
Average omnidirectional	59.7%	65.9%	68.2%
UltraZoom	59.3%	77.9%	38.0%
StereoZoom	59.0%	78.9%	31.5%
	Bilateral listeners		
	Signal location		
	Poorer ear	Front	Better ear
Processor Mic	49.7%	58.8%	59.6%
T-Mic 2	47.0%	59.8%	61.9%
P-Mic + T-Mic 2	46.6%	58.8%	63.7%
Average omnidirectional	47.8%	59.1%	61.7%
UltraZoom	22.4%	61.7%	35.3%
StereoZoom	24.7%	68.7%	30.4%

Abbreviations: CI, cochlear implant; HA, hearing aid.

Bold text represents mean scores for all 3 omnidirectional mic conditions.

mean bilateral CI speech recognition scores for speech directed to the better CI ear were 35.3 and 30.4% for UltraZoom and StereoZoom, respectively.

Mean bilateral CI speech recognition scores with *speech at 0°* were 58.8% for the processor microphone, 59.8% for the T-Mic 2, and 58.8% for the processor microphone + T-Mic 2. Beamforming resulted in greater speech recognition with speech at 0° as compared with all omnidirectional mics with mean scores of 61.7 and 68.7% correct with UltraZoom and StereoZoom, respectively.

Statistical Analysis: Omnidirectional Microphones Alone

Statistical analysis was first completed via a linear mixed model with the source azimuth (0°, 90°, and 270°), omnidirectional mic configuration (processor microphone, T-Mic 2, and processor microphone + T-Mic 2), and subject group (bimodal and bilateral) as the independent variables and TIMIT sentence recognition, in percent correct, as the dependent variable. This analysis was completed as a direct comparison to our previous paper with previous-generation technology.²²

We investigated main effects and all interaction terms for omnidirectional microphone locations alone. Statistical analysis revealed a statistically significant effect of source azimuth ($F_{(2, 82)} = 3.32, p = 0.04, \eta_p^2 = 0.07$), a significant effect of subject group ($F_{(1, 82)} = 6.66, p = 0.01, \eta_p^2 = 0.08$), no effect of mic configuration for omnidirectional conditions ($F_{(2, 82)} = 0.06, p = 0.94, \eta_p^2 = 0.001$), and no three-way interaction ($F_{(12, 82)} = 0.24, p = 0.99, \eta_p^2 = 0.03$). Because we found no main effect of mic configuration and no interaction effects, we collapsed across omnidirectional mic configuration for all subsequent analyses. Combined results are displayed in ► Fig. 2.

Statistical Analysis: Omnidirectional versus Beamforming

Statistical analysis investigating the effects of subject group, source azimuth, and mic type (omnidirectional vs. beamformer) as the independent variables on TIMIT sentence recognition in noise was completed using a linear mixed model. Recall that we averaged across omnidirectional mics (processor mic, T-Mic 2, and processor mic + T-Mic 2) given the lack of effect reported previously. In this analysis, we found a significant main effect of subject group ($F_{(1, 81)} = 16.83, p < 0.0001, \eta_p^2 = 0.17$), a significant main effect of source azimuth ($F_{(2, 81)} = 21.90, p < 0.0001, \eta_p^2 = 0.35$), a significant main effect of mic type ($F_{(2, 81)} = 3.75, p = 0.028, \eta_p^2 = 0.09$), a significant interaction between azimuth and mic type ($F_{(4, 81)} = 4.58, p = 0.002, \eta_p^2 = 0.18$), but no

other two- or three-way interactions were statistically significant.

Post hoc analyses were completed using all-pairwise multiple comparison using a Holm–Sidak statistic. For UltraZoom, there was a significant difference between scores obtained with speech presented to the better ear as compared with the front ($t = 4.4, p < 0.001, d = 2.3$) as well as for speech presented to the poorer ear as compared with the front ($t = 3.8, p < 0.001, d = 1.3$). There was no difference between scores obtained with speech from either the better or poorer hearing ear ($t = 0.58, p = 0.58, d = 0.2$). For StereoZoom, there was a significant difference between scores obtained with speech to the front as compared with the better ear ($t = 3.8, p < 0.001, d = 3.3$), speech presented to the front as compared with the poorer ear ($t = 6.0, p < 0.001, d = 1.7$), as well as between scores obtained with speech presented to the better versus poorer ear ($t = 2.2, p = 0.03, d = 0.6$). In this analysis for which we collapsed across all omnidirectional mic types, there were no significant differences between scores obtained with any of the source azimuths (front vs. better: $t = 0.19, p = 0.85, d = 0.2$; front vs. poorer: $t = 1.3, p = 0.35, d = 0.5$), poorer vs. better: $t = 1.5, p = 0.35, d = 0.6$).

Post hoc comparisons within a given source azimuth were also completed. For speech originating from the front, there were no statistically significant differences between any of the microphone types (omni vs. UltraZoom: $t = 0.99, p = 0.54, d = 0.7$; omni vs. StereoZoom: $t = 1.5, p = 0.35, d = 1.7$; UltraZoom vs. StereoZoom: $t = 0.53, p = 0.60, d = 1.0$); however, it

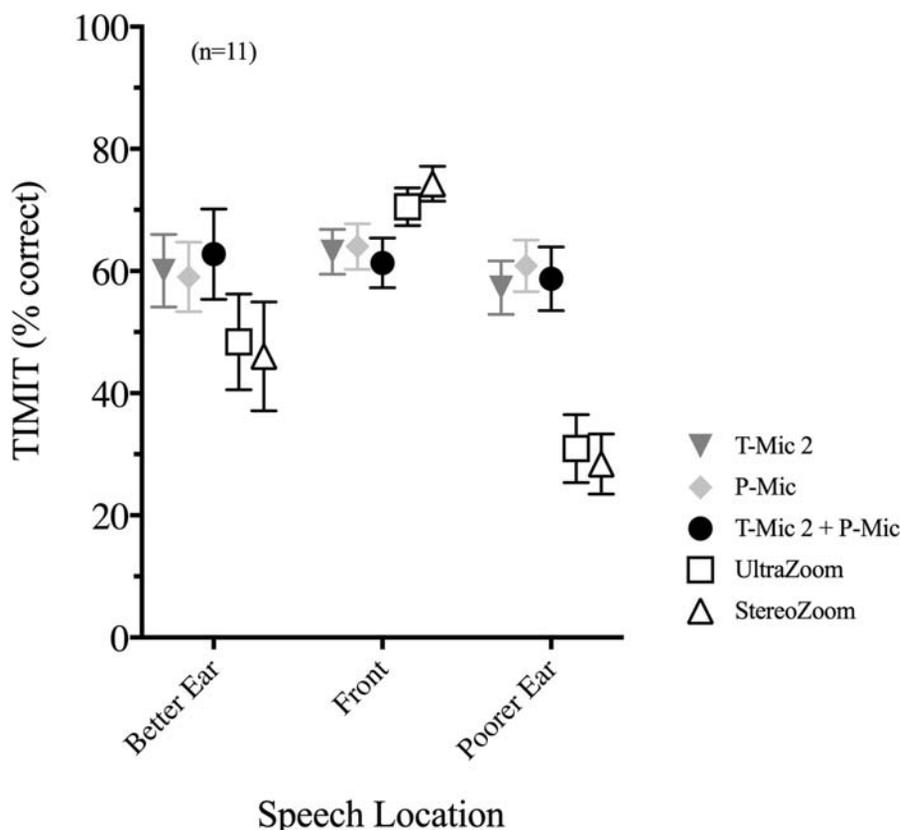


Fig. 2 Mean TIMIT sentence recognition (in percent correct) as a function of sound source azimuth. Error bars represent the standard error of the mean. TIMIT, Texas Instruments Massachusetts Institute of Technology.

should be noted that speech recognition at 0° was higher with UltraZoom (mean = 70.5%) and StereoZoom (mean = 78.9%) as compared with an omnidirectional microphone (mean = 62.83%), though this did not reach statistical significance. For speech originating from the poorer ear, there was a significant difference between scores obtained with omnidirectional and StereoZoom ($t = 3.2$, $p = 0.006$, $d = 0.4$), but no difference between omnidirectional and UltraZoom ($t = 1.5$, $p = 0.14$, $d = 0.4$) nor between UltraZoom and StereoZoom ($t = 1.7$, $p = 0.19$, $d = 0.03$). For speech originating from the better ear, there was a significant difference between scores obtained with omnidirectional and UltraZoom ($t = 3.6$, $p = 0.002$, $d = 2.0$), a significant difference between omnidirectional and StereoZoom ($t = 2.5$, $p = 0.03$, $d = 2.3$), but no significant difference between UltraZoom and StereoZoom ($t = 1.1$, $p = 0.29$, $d = 0.3$).

Aided Detection

Sound-field thresholds were measured using frequency-modulated warble tones at octave and interoctave frequencies from 250 to 6,000 Hz for 15 implanted devices. Stimuli were presented from a single loudspeaker at 0°. A two-way analysis of variance (ANOVA) was completed using microphone source and frequency as the independent variables and aided threshold (in dB hearing level) as the dependent variable; this dataset met the assumptions of normality and equal variance. Results indicated a significant effect of microphone ($F_{(1, 209)} = 38.68$, $p < 0.001$, $\eta_p^2 = 0.16$), significant effect of frequency ($F_{(6, 209)} = 6.20$, $p < 0.001$, $\eta_p^2 = 0.15$), and a significant interaction ($F_{(6, 209)} = 22.22$, $p < 0.001$, $\eta_p^2 = 0.39$). To summarize, aided detection was significantly lower (i.e., better) with the processor microphone when compared with the T-Mic 2 and this difference was dependent upon the frequency being tested. Post hoc analyses showed that the processor microphone yielded significantly better average aided thresholds for 2,000 Hz (21.7 vs. 24.7 dB; $t = 3.10$, $p = 0.003$; $d = 0.38$) and 6,000 Hz (18.0 vs. 30.3 dB; $t = 3.10$, $p = 0.003$; $d = 12.68$, $p < 0.001$; $d = 1.79$).

Microphone Output Results

To investigate significant differences in aided detection at some frequencies, physical measurements taken on a KEMAR mannequin are displayed in ►Fig. 3A,B. ►Fig. 3A plots the physical output of the processor microphone (gray line) and T-Mic 2 (black line) for a broadband, pink noise presented to the front of the listener. ►Fig. 3B shows the difference in the physical output level in dB, for a broadband, pink noise presented at 0° (dashed line) and at 90° azimuth for the processor microphone (gray line) and T-Mic 2 (black line). Here, a negative value indicates that the microphone output was higher for signals originating from 90° as compared with 0°. When averaged across the frequency range (86–9000 Hz), both the processor microphone and T-Mic 2 were more sensitive when the signal was presented from 90°, 2.1 dB and 6.0 dB, respectively. The overall output of the processor microphone was 6.9 dB greater than the T-Mic 2 at 0° (►Fig. 3A) and 3.0 dB greater at 90°. A two-way ANOVA using microphone source and azimuth as the independent variables

and output for pink noise as the dependent variables was completed encompassing the spectral range from 86 to 9,000 Hz; note however that the dataset did not meet assumptions of normality nor equal variance. Results indicated statistically significant main effects of mic ($F_{(1, 416)} = 21.14$, $p < 0.001$, $\eta_p^2 = 0.05$) and azimuth ($F_{(1, 416)} = 13.66$, $p < 0.001$, $\eta_p^2 = 0.03$), but no interaction ($F_{(1, 416)} = 3.4$, $p = 0.065$, $\eta_p^2 = 0.008$). Post hoc analyses showed that for the processor microphone, there was no difference between output for 0° and 90° ($t = 1.3$, $p = 0.19$, $d = 0.19$). For T-Mic 2, there was a statistically significant difference between 0° and 90° ($t = 3.92$, $p < 0.001$, $d = 0.51$). At 0°, there was a significant difference between the processor microphone and T-Mic 2 ($t = 4.56$, $p < 0.001$, $d = 0.66$). At 90°, there was no statistically significant difference between the processor microphone and T-Mic 2 ($t = 1.94$, $p = 0.053$, $d = 0.26$).

Discussion

Hypothesis 1: The T-Mic 2 would result in significantly higher speech recognition than all other omnidirectional microphone locations.

While the results of previous work suggest exclusive use of the T-Mic in patients with an AB device,^{19–22,32,33} we found that in the newest processor from AB, speech recognition performance was not significantly different across all omnidirectional microphones. This might be explained by the improved orientation of the processor microphone which now rests on the top of the pinna (rather than slightly behind) in the latest-generation sound processor. The physical devices also sit on the ear differently due to the design, size, and weight of the devices. Sensitivity differences between the microphones may also help to explain why we observed no significant differences in speech recognition between omnidirectional mic sources in this latest-generation device. For example, in the Harmony processor, both the processor microphone and the T-Mic are the same microphone part. In contrast, the T-Mic 2 and processor microphone in the Naída are two different microphones with two different frequency responses. Lastly, differing mechanical housings of these microphones may create frequency shaping. For example, the processor microphone in the previous-generation Harmony was more recessed than the current-generation Naída and it is well known that housing can block high-frequency signals that arrive off-axis. Despite the fact that we did not observe differences in speech recognition performance for the T-Mic 2 as compared with other omnidirectional microphone configurations (processor mic and processor mic + T-Mic 2), we recognize that there may be distinct advantages to a T-Mic 2 program. This study focused on speech recognition in noise with a roving source and a physical mic output. We did not investigate whether pinna effects afforded by T-Mic 2 placement may have resulted in sound-quality differences across mic configurations. Further, the T-Mic 2 allows for natural hand-held telephone placement as well as unobstructed use of circumaural headphones for listeners preferring this listening method over streaming.

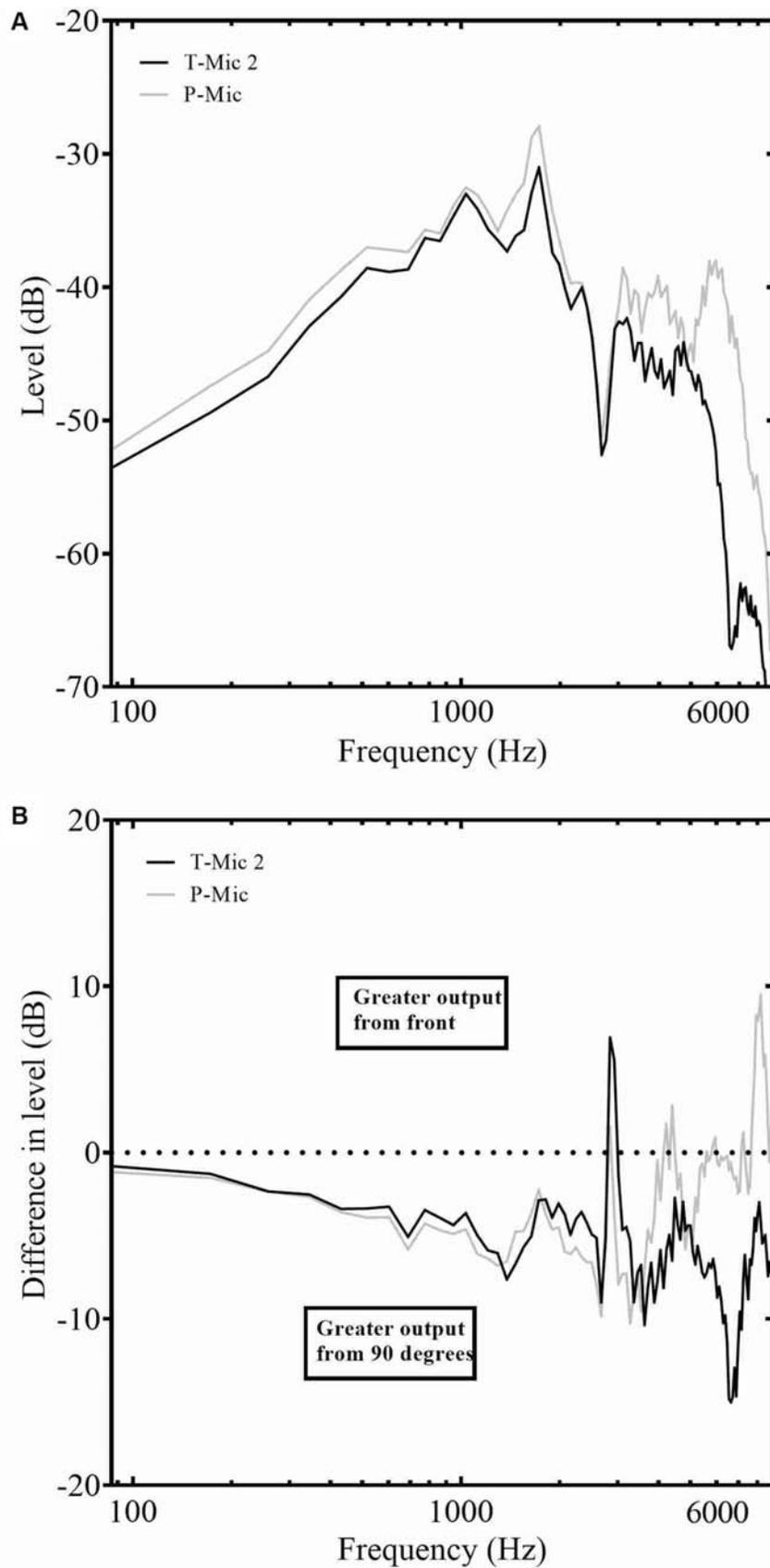


Fig. 3 (A) Overall output (in dB) when pink noise was presented from 0° azimuth for the T-Mic 2 (black line) and processor microphone (gray line) as a function of frequency. (B) The difference in the output of the T-Mic 2 (black line) and processor microphone (gray line) when pink noise was presented from 0° and then from 90° source azimuth as a function of frequency. Here, a negative value indicates that the microphone output was greater when the signal originated from 90° as opposed to 0°.

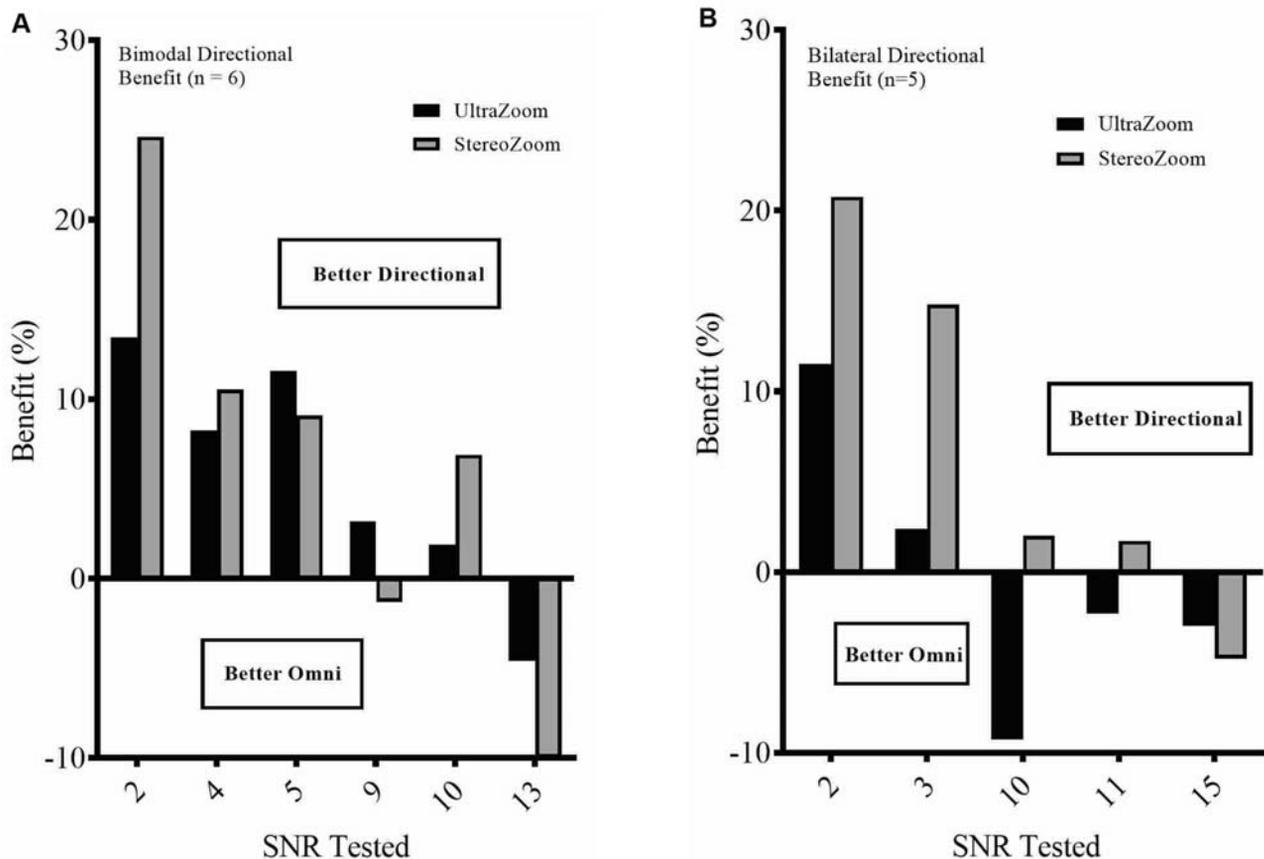


Fig. 4 (A and B) Individual benefit from UltraZoom (black) and StereoZoom (gray) over the best omnidirectional microphone performance for bimodal (panel A) and bilateral (panel B) listeners when speech was presented to the front of the listener. Note: participants were tested at one SNR (see procedures for SNR determination). SNR, signal-to-noise ratio.

Hypotheses 2 and 3: speech recognition would be greatest for omnidirectional microphones with speech originating from 90° and directional mics affording the best performance in speech originating from 0°.

Source azimuth did not significantly impact speech recognition in nondirectional microphone configurations. In directional programs, clinically significant directional benefit was observed in some individuals, but only at the most difficult SNRs (→ Fig. 4). These results are highly clinically relevant for CI programming audiologists. We can comfortably recommend any omnidirectional microphone setting, based on recipient preference. This is good news for audiologists and recipients as it allows for flexibility in choosing input source without sacrificing speech recognition. These results also inform the audiologist as to when they may consider a directional microphone program to suit a patient that seeks better speech recognition in noise without the use of an additional accessory. For example, we acknowledge that remote microphone systems would be a better choice for individuals who need a more favorable SNR to understand speech in noise than a directional mic can offer,⁶ but for individuals with good speech recognition in noise, a directional mic program is likely to help more than it would for individuals who do not have good recognition in noise.

Microphone Output

Similar to the work of Kolberg and colleagues,²² we did find differences in microphone output between omnidirectional microphones and differences within microphone output depending on signal azimuth. The current results are also in agreement with other publications that favor a side-source azimuth for BTE mics in HAs^{32,33} and in HRTF studies with CI processors^{19,20}. Interestingly, while all omnidirectional microphones in this study had greater output for signals at 90° versus 0°, this did not adversely affect speech recognition performance in bimodal or bilateral listeners. Microphone source, did, however, have a significant impact on aided detection at 2,000 and 6,000 Hz.

There are several reasons for this phenomenon: (1) the native physical response of the mics, (2) the location and/or orientation of the mics, and (3) the intensity of the stimulus. One additional explanation recently addressed in the literature is the presence of a partial head shadow, termed “face-shadow.”²³ Dwyer and colleagues explain the “face shadow” as the difference in speech recognition in unilateral hearing listeners when speech is presented to the front of the listener as compared with the CI side. They showed a significant correlation between the magnitude of the face-shadow effect and benefit with a contralateral routing of signal (CROS) device on the contralateral ear in both quiet speech (50 dB SPL) and speech in noise (+5 dB SNR). While

we did not see an impact of speech recognition as a function of source azimuth for any of the omnidirectional mics, any effect of source azimuth could have been overcome with the addition of the contralateral HA or CI, similar to the findings of Dwyer et al.²³ with the addition of the CROS device. Thus, future work may consider additional evaluation of the native frequency responses of specific microphones to isolate their contribution to differences in microphone output as microphone location is evaluated. For example, to determine if the T-Mic 2 does provide additional high-frequency gain due to pinna reflections, a recording could be made with the T-Mic 2 at the top of the pinna and another at the entrance of the canal.

While the broadband sensitivity of the processor microphone is higher, the T-Mic 2 did demonstrate higher output from ~2,000–3,000 and 4,000–5,000 Hz. This higher output with T-Mic 2 is likely due to placement at the opening of the ear canal. Additionally, the contribution/detriment to the frequency components of the signal as it reaches the user at different source azimuths could be controlled further by investigating them in the free field.

Study Limitations

It is important to recognize the limitations of the current study. The first limitation of this study was the sample size ($n = 11$). Further investigation with larger sample sizes is warranted. Second, only recipients of one manufacturer's device were used. However, AB is currently the only manufacturer that has both an integrated processor microphone and a microphone that rests at the entrance of the canal in the same external sound processor. The third consideration is that the influence of different ear shapes, sizes, and orientations could also account for differences in T-Mic 2 output, an effect that was not accounted for in the current study. Additionally, microphone recordings were only captured from one Naída CI Q90 processor—there is the potential for slight variability if these same measures were recorded from a sample of multiple processors. Fourth and finally, recordings from a single microphone at different locations were not made—doing so would have allowed us to control for the impact of the microphone-frequency response and isolate the acoustic differences that arise from microphone location. With consideration to the lack of differences observed in speech recognition, we did not feel further analysis of the microphone output added anything to the primary aims of the current study, which was to investigate the impact of microphone location and signal azimuth on speech recognition.

Conclusion

Our hypotheses were motivated by analysis of previous literature that examined the effects of microphone location in CIs and in HAs. Thus, we expected that T-Mic 2 would yield the best speech recognition in our study cohort and that off-axis signal presentation would yield the greatest speech recognition. We also expected to see a significant benefit in noise when using beamforming, with binaural beamforming (i.e., StereoZoom) offering the best speech recognition in noise when speech originates from

the front of the user. Our main study findings were as follows:

- Unlike previous-generation sound processors, we found no statistically significant effect of omnidirectional microphone location on speech recognition in noise from any source azimuth.
- While mean speech recognition for UltraZoom and StereoZoom was greater than for omnidirectional microphone modes, these differences were not statistically significant.
- Directional microphones proved effective, but only at the most difficult SNRs.

Our secondary findings were as follows:

- All omnidirectional microphones offered better speech recognition over beamforming when the signal is presented off-axis (i.e., 90°, 270°).
- Omnidirectional microphone output is greatest when the signal is presented off-axis.
- The processor microphone output was greater than the T-Mic 2 when the signal is presented to the front of the listener, which contributed to better aided detection at 2 and 6 kHz.

Funding

Funding was provided by both NIDCD R01 DC009404 (investigator effort) and AB (participant remuneration).

Conflict of Interest

R.H.G. reports a grant from the NIH NIDCD (DC009404) and a grant from Advanced Bionics to Vanderbilt University Medical Center, during the conduct of the study. R.H.G. is on the audiology advisory board for Advanced Bionics, LLC and Cochlear Americas and the clinical advisory board for Frequency Therapeutics. All other authors report no conflict of interest.

Acknowledgments

Portions of these data were presented at the American Auditory Society meeting in Scottsdale, AZ on March 3, 2017.

References

- 1 Gifford RH, Dorman MF, McKarns SA, Spahr AJ. Combined electric and contralateral acoustic hearing: word and sentence recognition with bimodal hearing. *J Speech Lang Hear Res* 2007;50(04):835–843
- 2 Dorman MF, Gifford RH, Spahr AJ, McKarns SA. The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiol Neurotol* 2008;13(02):105–112
- 3 Neuman AC, Svirsky MA. Effect of hearing aid bandwidth on speech recognition performance of listeners using a cochlear implant and contralateral hearing aid (bimodal hearing). *Ear Hear* 2013;34(05):553–561
- 4 Neuman AC, Zeman A, Neukam J, Wang B, Svirsky MA. The effect of hearing aid bandwidth and configuration of hearing loss on bimodal speech recognition in cochlear implant users. *Ear Hear* 2019;40(03):621–635
- 5 Gifford RH, Dorman MF, Sheffield SW, Teece K, Olund AP. Availability of binaural cues for bilateral implant recipients and

- bimodal listeners with and without preserved hearing in the implanted ear. *Audiol Neurotol* 2014;19(01):57–71
- 6 Dorman MF, Gifford RH. Speech understanding in complex listening environments by listeners fit with cochlear implants. *J Speech Lang Hear Res* 2017;60(10):3019–3026
 - 7 Yawn RJ, O'Connell BP, Dwyer RT, et al. Bilateral cochlear implantation versus bimodal hearing in patients with functional residual hearing: a within-subjects comparison of audiologic performance and quality of life. *Otol Neurotol* 2018;39(04):422–427
 - 8 Holder JT, Levin LM, Gifford RH. Speech recognition in noise for adults with normal hearing: age-normative performance for AzBio, BKB-SIN, and QuickSIN. *Otol Neurotol* 2018;39(10):e972–e978
 - 9 Holder JT, Sheffield SW, Gifford RH. Speech understanding in children with normal hearing: sound field normative data for BabyBio, BKB-SIN, and QuickSIN. *Otol Neurotol* 2016;37(02):e50–e55
 - 10 Dorman MF, Natale S, Spahr A, Castioni E. Speech understanding in noise by patients with cochlear implants using a monaural adaptive beamformer. *J Speech Lang Hear Res* 2017;60(08):2360–2363
 - 11 Spahr AJ, Dorman MF, Litvak LM, et al. Development and validation of the pediatric AzBio sentence lists. *Ear Hear* 2014;35(04):418–422
 - 12 Mosnier I, Mathias N, Flament J, et al. Benefit of the UltraZoom beamforming technology in noise in cochlear implant users. *Eur Arch Otorhinolaryngol* 2017;274(09):3335–3342
 - 13 Jansen S, Luts H, Wagener KC, et al. Comparison of three types of French speech-in-noise tests: a multi-center study. *Int J Audiol* 2012;51(03):164–173
 - 14 Holder JT, Taylor AL, Sunderhaus LW, Gifford RH. Effect of microphone location and beamforming technology on speech recognition in pediatric cochlear implant recipients. *J Am Acad Audiol* 2020. doi: 10.3766/jaaa.19025
 - 15 Buechner A, Dyballa KH, Hehrmann P, Fredelake S, Lenarz T. Advanced beamformers for cochlear implant users: acute measurement of speech perception in challenging listening conditions. *PLoS One* 2014;9(04):e95542
 - 16 Ernst A, Anton K, Brendel M, Battmer RD. Benefit of directional microphones for unilateral, bilateral and bimodal cochlear implant users. *Cochlear Implants Int* 2019;20(03):147–157
 - 17 Vroegop JL, Homans NC, Goedegebure A, Dingemans JG, van Immerzeel T, van der Schroeff MP. The effect of binaural beamforming technology on speech intelligibility in bimodal cochlear implant recipients. *Audiol Neurotol* 2018;23(01):32–38
 - 18 Pearsons KS, Bennett RL, Fidell S. *Speech levels in various noise environments*. Washington, DC: Office of Health and Ecological Effects, Office of Research and Development, US EPA; 1977
 - 19 Mantokoudis G, Kompis M, Vischer M, Häusler R, Caversaccio M, Senn P. In-the-canal versus behind-the-ear microphones improve spatial discrimination on the side of the head in bilateral cochlear implant users. *Otol Neurotol* 2011;32(01):1–6
 - 20 Aronoff JM, Freed DJ, Fisher LM, Pal I, Soli SD. The effect of different cochlear implant microphones on acoustic hearing individuals' binaural benefits for speech perception in noise. *Ear Hear* 2011;32(04):468–484
 - 21 Gifford RH, Revit LJ. Speech perception for adult cochlear implant recipients in a realistic background noise: effectiveness of pre-processing strategies and external options for improving speech recognition in noise. *J Am Acad Audiol* 2010;21(07):441–451, quiz 487–488
 - 22 Kolberg ER, Sheffield SW, Davis TJ, Sunderhaus LW, Gifford RH. Cochlear implant microphone location affects speech recognition in diffuse noise. *J Am Acad Audiol* 2015;26(01):51–58, quiz 109–110
 - 23 Dwyer RT, Kessler D, Butera IM, Gifford RH. Contralateral routing of signal yields significant speech in noise benefit for unilateral cochlear implant recipients. *J Am Acad Audiol* 2019;30(03):235–242
 - 24 Keidser G, Dillon H, Flax M, Ching T, Brewer S. The NAL-NL2 prescription procedure. *Audiology Res* 2011;1(01):e24
 - 25 Revit LJ, Killion MC, Compton-Conley CL. Developing and testing a laboratory sound system that yields accurate real-world results. *Hear Rev* 2007;14(11):54–62
 - 26 Compton-Conley CL, Neuman AC, Killion MC, Levitt H. Performance of directional microphones for hearing aids: real-world versus simulation. *J Am Acad Audiol* 2004;15(06):440–455
 - 27 Lamel L, Kassel RH, Seneff S. Speech database development: design and analysis of the acoustic-phonetic corpus. *Proceedings of DARPA Speech Recognition Workshop*. 1989:100–109
 - 28 Loizou PC, Dorman M, Poroy O, Spahr T. Speech recognition by normal-hearing and cochlear implant listeners as a function of intensity resolution. *J Acoust Soc Am* 2000;108(5, Pt 1):2377–2387
 - 29 Dorman MF, Loizou PC, Spahr AJ, Dana CJ. Simulations of combined acoustic/electric hearing. *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology*. 2003:199–201
 - 30 Dorman MF, Spahr AJ, Loizou PC, Dana CJ, Schmidt JS. Acoustic simulations of combined electric and acoustic hearing (EAS). *Ear Hear* 2005;26(04):371–380
 - 31 King SE, Firszt JB, Reeder RM, Holden LK, Strube M. Evaluation of TIMIT sentence list equivalency with adult cochlear implant recipients. *J Am Acad Audiol* 2012;23(05):313–331
 - 32 Festen JM, Plomp R. Speech-reception threshold in noise with one and two hearing aids. *J Acoust Soc Am* 1986;79(02):465–471
 - 33 Pumford JM, Seewald RC, Scollie SD, Jenstad LM. Speech recognition with in-the-ear and behind-the-ear dual-microphone hearing instruments. *J Am Acad Audiol* 2000;11(01):23–35
 - 34 Spahr AJ, Dorman MF, Litvak LM, et al. Development and validation of the AzBio sentence lists. *Ear Hear* 2012;33(01):112–117