Trends in Hearing
Volume 22: 1–12

DOI: 10.1177/2331216518804945
journals.sagepub.com/home/tia

Effect of Hearing Aid Directionality and Remote Microphone on Speech Intelligibility in Complex Listening Situations

Kirsten C. Wagener[1,2], Matthias Vormann[1,2], Matthias Latzel[3], and Hans E. Mülder[4]

Abstract

Remote microphones (RMs) have been developed to support hearing aid (HA) users in understanding distant talkers. In traditional clinical applications, a drawback of these systems is the deteriorated speech intelligibility in the near field. This study investigates advantages and disadvantages of clinical RM usage and the effects of different directionality settings of the HAs in complex listening situations in the laboratory. Speech intelligibility was investigated in 15 experienced severely hearing impaired participants in a noisy environment using a dual-task test paradigm where the tasks were presented from either a near field or a far field loudspeaker. Primary and secondary tasks were presented simultaneously so attention had to be shared on both tasks. In a second experiment, two speech intelligibility tests were presented from either the near field or the far field loudspeaker. The tests were interleaved to simulate a complex listening situation with shifting attention. Directional HA microphones yielded better performance than omnidirectional microphones (both combined with a RM) in near field when analyzing both tasks of the dual-task experiment separately. Furthermore, the integrated dual-task test results showed better performance with directional HA microphones compared with the omnidirectional setting (both cases in combination with a RM). These findings were confirmed by the results of the interleaved speech intelligibility test.

Keywords

hearing loss, hearing aids, directionality, remote microphone, dual-task speech intelligibility

Date received: 31 January 2018; revised: 7 September 2018; accepted: 11 September 2018

Introduction

One of the most common problems for individuals with hearing loss is understanding speech in complex listening environments. Listening is often difficult when there is excessive background noise, reverberation, or a large distance between the source of interest and the individual with hearing loss that would minimize helpful modulations in the perceived signal (Bronkhorst, 2000; International Electrotechnical Commission, 2011; Plomp, 1977). To overcome these three main factors contributing to adverse listening situations, individuals with hearing loss require better signal-to-noise ratios (SNRs) than those with normal hearing: Wilson, McArdle, and Smith (2007) reported that individuals with a moderate hearing loss require an increased SNR of up to 10 dB to achieve the same speech understanding as individuals with normal hearing. For individuals with a severe-to-profound hearing loss, an average SNR increase of up to 20 dB is required (Killion & Niquette, 2000). The only hearing aid (HA) feature that has been shown to significantly increase speech understanding in noise is the directional microphone (Dillon, 2012; Killion, 2004).

Most directional microphones provide a 4 to 5 dB SNR improvement on average, which offers a substantial

[1]Hörzentrum Oldenburg GmbH, Germany
[2]Cluster of Excellence Hearing4all, Oldenburg, Germany
[3]Sonova AG, Stäfa, Switzerland
[4]Phonak Communications, Murten, Switzerland

Corresponding author:
Kirsten C. Wagener, Hörzentrum Oldenburg GmbH, Marie-Curie-Str. 2, D-26129 Oldenburg, Germany.
Email: k.wagener@hoerzentrum-oldenburg.de

benefit to HA users and this is known to contribute significantly to personal device satisfaction (Kochkin, 1996).

But there are still some limitations, for example, a 4 to 5 dB SNR improvement is not enough to address the average 10 to 20 dB SNR improvement needed for individuals with moderate to severe hearing loss. In addition, directional microphones are intended for use in the near field or in approximately 1.5 m distance from the source of interest. Blazer (2007) reported that students with hearing loss were able to achieve 95% on speech recognition tasks when they were 1.8 m apart from the source of interest and only 60% when they were 7.3 m apart from the source of interest. Thus, directional microphones provide great benefit in environments with low reverberation, and when the source of interest is positioned in front of the HA wearer (Kim & Kim, 2014).

Individuals who need additional SNR improvement beyond the performance of traditional directional microphones can use remote microphones (RMs). RMs are used to reduce background noise, decrease reverberation, and shorten the distance between the source of interest and the individual with hearing loss (Johnson & Seaton, 2011). RMs are intended for far field use and have historically been accomplished with frequency modulation transmission (FM). In addition to the SNR improvement of RM by picking up the target sound at a close distance, a traditional analogue FM system provides a fixed FM advantage level independent of the background noise. This is usually seen as a +10 dB increase in decibel output level when the FM signal is added to the signal from the HA's own microphone. The FM advantage of 10 dB is a default FM system parameter value that is in line with the professional consensus which states that "the FM system should increase the level of the perceived speech, in the listener's ear, by at least 10 dB relative to reception by hearing aid only" (American Academy of Audiology Clinical Practice Guidelines, 2011). The purpose of the additional 10 dB is to emphasize the voice of a teacher in a class in advantage to the peers' voices in class, for example. FM systems have shown significant benefit for both HA users (Anderson & Goldstein, 2004; Thibodeau, 2014) and cochlear implant (CI) users (Wolfe et al., 2009, 2013). Digital wireless systems are able to provide an additional SNR improvement compared with analogue FM systems. Wolfe et al. (2013) found that CI users performed better with adaptive digital wireless systems compared with analogue and fixed-gain FM systems in high levels of noise. With CI users, the general benefit from digital RM was confirmed in recent studies by Wolfe et al. (2015) and De Ceulaer et al. (2016) or Vroegop, Dingemanse, Homans, and Goedegebure (2017) for bimodal CI users. Further, Thibodeau (2014) found that adaptive digital wireless RM technology resulted in significantly better speech recognition in loud noise compared with (a) normal-hearing listeners without RM, (b) analogue fixed-gain FM RM, and (c) analogue adaptive FM RM. This is partly due to the additional automatic receiver gain amplification of 1 to 30 dB in noisy environments (Johnson & Seaton, 2011; Thibodeau, 2014).

Hence, in clinical applications RM are used to support hearing instrument users to understand distant talkers in complex acoustic environments (i.e., the distant talker wears the RM). By this use, RM often deteriorates signal perception and thus speech intelligibility in the near field since the RM signal is increased in output level compared with the HA microphone output. This is different to the usage of external microphones in research about acoustic sensor networks (Bertrand & Moonen, 2009; Gößling, Marquardt, & Doclo, 2017; Szurley, Bertrand, Van Dijk, & Moonen, 2016). Therefore, listeners may usually have problems to follow a speech at a large family celebration via the RM while simultaneously getting comments from the table neighbors. In contrast to the literature about external microphones in acoustic sensor networks, which only includes simulations of the external microphone effect, the present study presents speech intelligibility measurements with hearing impaired listeners using clinically available hearing devices and RMs. To mimic a family celebration conversation situation in the laboratory, a dual-task speech intelligibility test paradigm was developed where the primary and secondary tasks were simultaneously presented from two loudspeakers in a diffuse restaurant noise environment. The primary task was presented either from the loudspeaker in the near field (1.4 m distance from listener; the "neighbor at the opposite side of the table") or a loudspeaker in the far field (6.4 m distance from listener; "the speaker"). The secondary task was presented simultaneously from the other loudspeaker.

Since it is also common for a listener to change who and what they want to listen to in a given situation, an interleaved speech intelligibility test with alternating near field and far field presentations was performed.

In the past, digital HAs were provided with two analogue to digital converters (AD). When using a RM, one of the AD was required to utilize the RM, leaving only one AD for further processing tasks of the HA. Therefore, only a single microphone mode (omnidirectional) was possible. More recently, HAs are available and were used in the present study, which utilize three analogue to digital converters in the input stage of the hearing device. This allows for both microphones of the hearing device to be used to build a subtractive directional microphone system (Dillon, 2012) while also using the RM. The benefit of RM technology in combination with omnidirectional HA microphones versus directional

HA microphones in children was investigated by Jones and Rakita (2016). They found better speech intelligibility in noise up to 25% with directional HA settings compared with omnidirectional settings. This study shows the performance of RM with omnidirectional versus directional HA microphones in either near field or far field target signals. However, to date, no studies have investigated the effect of the source of interest and therefore attention shifting between the near field and far field or simultaneous shared attention.

The present study is therefore unique in aiming to simulate a realistic listening environment where the source of interest shifts from being close to the HA wearer to being further apart, or the listener wishes to divide the attention. The effect of a remote adaptive digital wireless microphone as well as different directionality settings of the HAs were investigated with the target group of RM technology, severely hearing impaired participants who are experienced HA users. The hypotheses followed in this study were as follows: (a) Speech intelligibility in near field is higher with directional HA microphones compared with omnidirectional. (b) Speech intelligibility in far field is not affected by the directionality of HAs. (c) Effects are expected to be larger in listening situations with shifting attention compared with situations with shared attention.

Materials and Methods

Ethical approval for all experimental procedures was obtained from the ethics committee of the University of Oldenburg (reference number Drs. 36/2015). Prior to any data collection, written informed consent was obtained from all participants. Participants were paid on an hourly basis for their participation.

Participants

Fifteen (4 women and 11 men) severely hearing impaired listeners participated in the measurements. All were experienced HA users for more than 9 years but had no experience with RM in combination with their hearing devices. Age ranged from 63 to 83 years with a mean age of 72.3 (±6.4) years. Figure 1 shows the mean pure-tone audiogram thresholds of the participants.

In the run-up to the laboratory measurements, the participants completed a training session of the dual-task test and the single tasks, respectively. Participants, who were mentally overcharged by either one of the tasks or showed very low performance, were excluded from the following measurements. Fifteen of 20 invited listeners passed the criteria to participate in the main study. The participants were provided with the test devices during the study period and returned them afterwards.

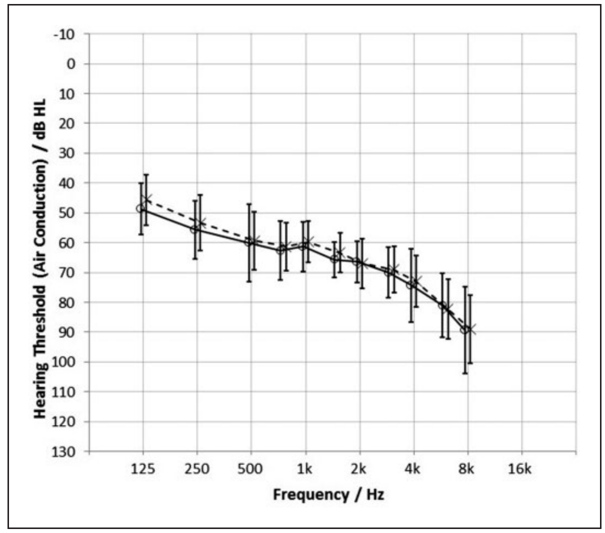

Figure 1. Average hearing loss (air conduction) of participants: mean pure-tone audiogram thresholds with standard deviation for left (solid line) and right (dashed line) ears.

Physical Test Setup

All measurements were carried out in a sound treated seminar room with an average size of 13.5 m × 7.2 m, room height of 3.4 m, and a critical distance for an omni-source loudspeaker of 1.5 m. The reverberation time (mean over the frequency range from 31.5 Hz to 16 kHz) of this room is 0.33 s. See Figure 2 for a schematic top view of the measurement room.

Eight Genelec 8030A Studio Monitors were placed at a distance of 1.5 m along the two longer walls of the room and radiating outwards (against the walls) to improve the diffusiveness of the presented background noise, which was a multimicrophone recording of restaurant noise at a level of 62 dB SPL, measured at the position of the participants head. A noise presentation level >60 dB SPL was chosen in order for it to be audible and result in an auto selection of the directional microphone. The switching algorithm would expect a noise level of more than 50 dB SPL and a portion of minimum 80% of "Speech in Noise" class to be classified in order to switch into the directional microphone mode. Nevertheless, all tests in the laboratory were performed with fixed directional microphone settings and manual switching of HA conditions. The target speech signals were presented via Tannoy System 800 A HR loudspeakers, placed at a distance of 1.4 m (near field) and 6.4 m (far field) to their heads. The near field loudspeaker was placed *inside* the critical distance of the room, meaning that the directional sound was stronger than the diffuse sound (sum of reflections) at the position of the subject. On the other hand, the far field loudspeaker was placed *outside* the critical distance.

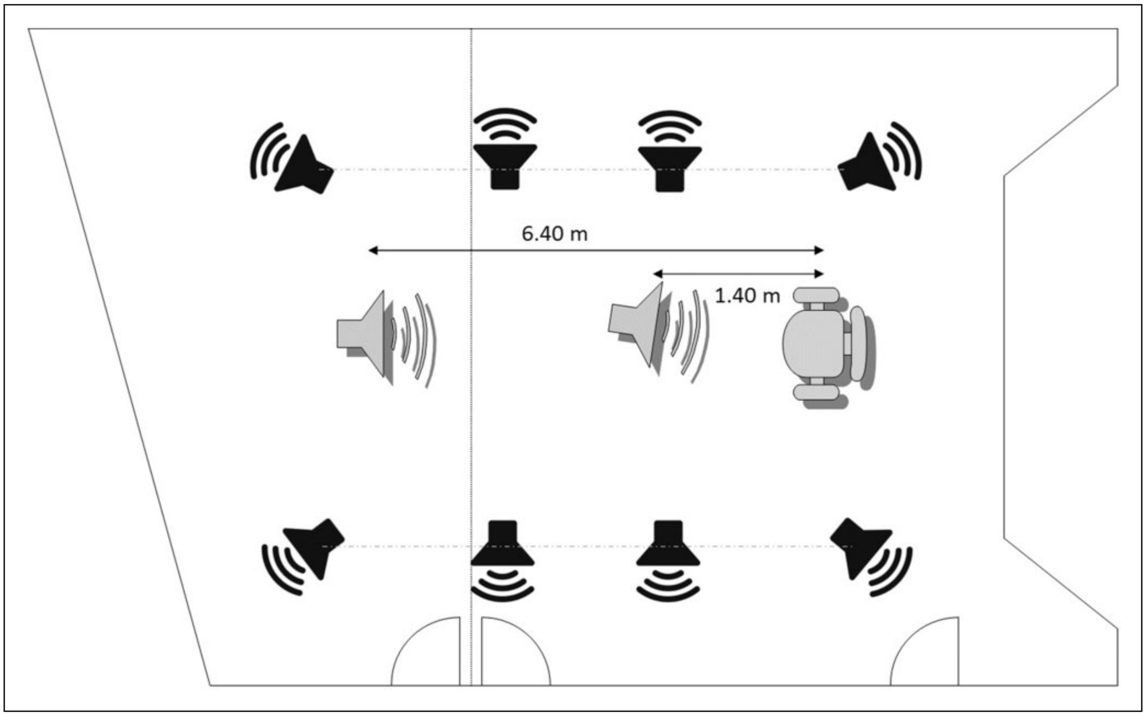

Figure 2. Schematic top view of the measurement room. Participant is seated on the chair in the lower middle. Near field speaker is positioned 1.4 m in front of the subject, far field speaker 6.4 m away from the subject. At the outside of the room, eight speakers are used to produce the background noise (restaurant noise).

These loudspeakers included a coaxial chassis and were therefore close to the emitting characteristics of a speaker's mouth. The RM was mounted in front of the far field speaker, simulating the RM position for a human speaker wearing a neck loop (20 cm in front of the coaxial chassis and 20 cm below the acoustic axis, see Figure 3).

As measurement software the Oldenburg measurement applications (Hörtech gGmbH Oldenburg, Germany) as well as in-house developed scripts implemented in MATLAB (MathWorks, Natick, USA) were used. The software was run on a personal computer located inside a control room next door. A RME (Haimhausen, Germany) HDSP 9652 soundcard was used, which was connected symmetrically to the loudspeakers via two RME (Haimhausen, Germany) ADI-8 Pro AD/DA converters.

For the dual-task speech intelligibility test, a tablet computer was used by the participants as the entering device for the primary task. It was connected via Intranet to the measurement personal computer.

Figure 3. Position of the RM in front of the coaxial chassis of the far field speaker. The folding yardstick is used to show the spacing between far field speaker and RM—it is removed during the tests.

Speech Stimuli, Tasks, and Measures

Dual-task speech intelligibility measurements. The primary task used the speech material of the Oldenburg sentence test (OLSA; Wagener & Brand, 2005; Wagener, Brand, Kühnel, & Kollmeier, 1999) with a female speaker

(Wagener, Hochmuth, Ahrlich, Zokoll, & Kollmeier, 2014). The OLSA is the German version of a Matrix test with the fixed grammatical structure: *name verb numeral adjective object* (Kollmeier et al., 2015). The OLSA material was not used as a normal speech intelligibility test in this study where the listener would repeat or mark all words that were understood. The sentences were successively presented with an inter-sentence pause of 1 s at a fixed presentation level of 65 dB SPL (near field) and 70 dB SPL (far field). These presentation levels were chosen according to pre-tests to allow for appropriate task execution. The number of presented sentences was variable, corresponding to the length of the secondary task measurement. The task of the participant was to consecutively recognize only the name in the OLSA sentences. As soon as the name was perceived, participants were tasked to select it from a list of the 10 possible names of the entire OLSA test, which was presented on the tablet screen. The performance measure of the primary task was the *correct response rate* of the presented names calculated as a percentage score. The task of the primary task was chosen to mimic the process of following a speech or conversation by not getting precisely each word correct but to follow at least part of the speech/conversation over time.

The secondary task was a standard sentence intelligibility test with every day sentences, the Göttingen sentence test (GÖSA; Kollmeier & Wesselkamp, 1997). The sentences were spoken by a male speaker. The task of the participant was to verbally repeat the perceived words. Lists of 20 sentences were used throughout the measurements; lists of 12 sentences were used during training. Since the test lists of GÖSA should not be used twice in short succession due to sentence familiarity, specific training lists were used for the dual-task training. Therefore, during all measurements, each particular GÖSA test list was only used once per subject. The sentences were presented at an individually chosen, fixed presentation level that was determined at the end of the training sequence, see section Measurements. This presentation level was fixed for all measurement conditions. The performance measure of the secondary task was percentage *speech intelligibility* based on word scoring.

Both tasks were performed simultaneously during the dual-task measurements. In the default automatic program of the included HAs, the directional microphone setting "faded in" in approximately 15 to 20 s (time needed for situation classification and smoothly switching to directional mode). Although no automatic programs were used throughout the study (only manual switching of directionality was used), the first speech presentation for both tasks started 20 s after the restaurant noise had started in order to generate stable and reliable test conditions.

Interleaved speech intelligibility measurements. The interleaved speech intelligibility measurements were set up from two individual OLSA tests with a male speaker: One test list was presented via the near field speaker; the other OLSA list was presented via the far field speaker. Both tests were carried out with 20 sentences each. "Interleaved" in this context means, that the two tests were run nested, that is, the speaker for the next presentation was chosen randomly. Therefore, the participant did not know if the next sentence would be presented from far field or near field. The sentences were presented at an individually chosen, fixed presentation level that was determined at the end of the training sequence, see section Measurements.

The task of the participant was to repeat all perceived words. The performance measure of this task was the percentage speech intelligibility for the measurement in the near field and far field, separately. Word scoring was used to count the correct answers of the subject.

Screening

The dual task is rather challenging, especially for severely hearing impaired listeners. Therefore, all listeners were first required to pass training and screening with their own hearing instruments. On that basis, those listeners who were not capable of executing the dual-task paradigm sufficiently were excluded from the trial. Table 1 summarizes the measurements performed with the own hearing instruments. All measurements were performed via the near field loudspeaker in fixed order from top to bottom.

First of all, the GÖSA was measured in Goenoise (speech-shaped noise of GÖSA; Kollmeier & Wesselkamp, 1997) at a fixed level of 65 dB SPL with adaptively controlled speech level to find the individual speech reception threshold (SRT). Next GÖSA was measured adaptively in restaurant noise at a fixed noise presentation level of 62 dB SPL. "Adaptive SRT_{50}" represents the adaptive procedure that converges to 50% intelligibility (Brand & Kollmeier, 2002), and "adaptive SRT_{80}" represents the adaptive procedure that converges to 80% intelligibility.

The speech presentation level of the primary task (OLSA) was fixed at 65 dB SPL. The task was measured in quiet and in restaurant noise. The idea was to get used to the OLSA sentences and to the procedure of specifying the names on the tablet.

After carrying out the primary and secondary tasks separately, three measurements of the dual task were undertaken for introduction and screening.

To participate in the main study, the listeners had to be able to manage all tasks reliably and the SRT_{50} difference of GÖSA in dual task (Measurement 7) compared with secondary task alone (Measurement 2) had to be less than

Table 1. Training and Screening Measurements for the Dual-Task Paradigm With Own Hearing Instruments.

	Task	Speech material	Measure	Spatial setup	Interfering noise	Number of sentences
1	Secondary alone	GÖSA	Adaptive SRT_{80}	S_0N_0	Goenoise 65 dB	12
2	Secondary alone	GÖSA	Adaptive SRT_{50}	S_0N_{diff}	Restaurant noise 62 dB	12
3	Primary alone	OLSA	Correct response rate	S_0	Quiet	20
4	Primary alone	OLSA	Correct response rate	S_0N_{diff}	Restaurant noise 62 dB	20
5	Dual	OLSA/GÖSA	Correct response rate/adaptive SRT_{50}	S_0N_{diff}	Restaurant noise 62 dB	GÖSA 12/OLSA variable
6	Dual	OLSA/GÖSA	Correct response rate/adaptive SRT_{50}	S_0N_{diff}	Restaurant noise 62 dB	GÖSA 12/OLSA variable
7	Dual	OLSA/GÖSA	Correct response rate/adaptive SRT_{50}	S_0N_{diff}	Restaurant noise 62 dB	GÖSA 20/OLSA variable

Note. All measurements were performed via the near field loudspeaker. The spatial setup is labeled in the following way: The direction of the Speech "S" and the Noise "N" indicated as subscript, is given as the angle relative to the front of the participants. The subscript "diff" means a diffuse noise. GÖSA = Göttingen sentence test; OLSA = Oldenburg sentence test; SRT = speech reception threshold.

5 dB, while the correct response rate of OLSA in the dual task (Measurement 7) had to exceed 25%.

Five of the 20 listeners did not fulfil these criteria and therefore only 15 listeners participated in the main study.

HA Conditions

After the screening, the 15 participants were acclimatized to the test devices in a daily life usage period of about 4 weeks including fine tuning after 1 week. If needed, additional fine tuning was possible according to the needs of the participants. As test devices Phonak Naida V90 SP were used coupled with individual ear molds, applying the manufacturer's own prescription rule. No RM was used during the first daily life period.

Prior to the laboratory measurements in Visit 1, Phonak Roger 18 receivers and Roger Pens were connected to the HAs and used as RM throughout the measurements. No acclimatization to the RM devices was obtained at that time. Thereafter, the RM was used by the participants during the second daily life period.

In the laboratory tests, two different HA conditions were investigated: (a) omnidirectional HA microphones with remote microphone (Omni + RM) and (b) directional HA microphones with remote microphone (Dir + RM). In addition, the binaural beamformer of the HAs with focused front directionality was tested without RM usage. These data are not reported in the present article. All measurements in the laboratory were performed with fixed HA microphone directionality settings. The directivity index (DI) of the HA measured according to ANSI S3.35-2010 standard (American National Standard, 2010) can be seen in Figure 4 for omnidirectional and directional microphone settings. The SII-DI (average speech weighting according to ANSI S3-35, Table C1, column 1) of the directional microphone setting has been calculated to 4.5 dB.

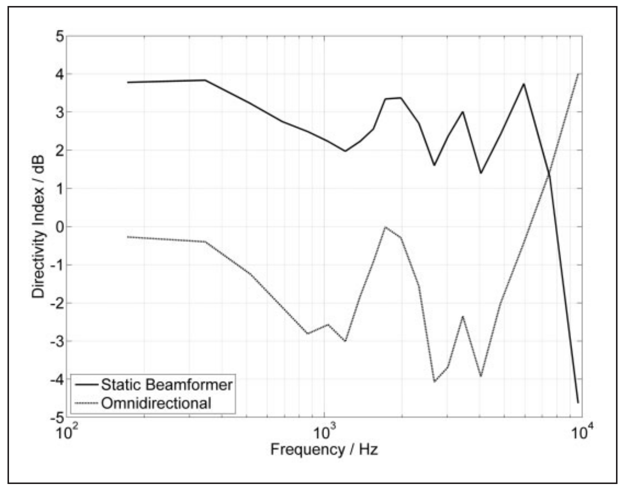

Figure 4. DI-2D of the test devices for omnidirectional (dotted line) and directional mode (solid line). The measurements were performed based on ANSI-S3-35: 2010 (ANSI noise, 75 dB SPL).

Measurements

After acclimatization to the test devices, the laboratory measurements were distributed across three to four visits with a maximum duration of 2 h each. Two participants needed four visits instead of three due to longer measurement time. These participants performed the measurement block of the interleaved speech intelligibility tests in the additional visit.

In the first visit, additional training of the dual-task test was performed with the test devices in the focused front directionality setting of the binaural beamformer but without RM, applying sound presentation via the near field loudspeaker without RM (Table 2).

The secondary task presentation level for all following dual-task measurements was defined by adding 3 dB to the SRT_{50} of the last training measurement (No. 5 in Table 2),

Table 2. Training Measurements for the Dual-Task Paradigm With Test Devices at Visit 1.

	TASK	Speech material	Measure	Spatial setup	Number of sentences
1	Secondary alone	GÖSA	Adaptive SRT_{50}	S_0N_{diff}	12
2	Primary alone	OLSA	Correct response rate	S_0N_{diff}	20
3	Dual	OLSA/GÖSA	Correct response rate/adaptive SRT_{50}	S_0N_{diff}	GÖSA 12/OLSA variable
4	Dual	OLSA/GÖSA	Correct response rate/adaptive SRT_{50}	S_0N_{diff}	GÖSA 12/OLSA variable
5	Dual	OLSA/GÖSA	Correct response rate/adaptive SRT_{50}	S_0N_{diff}	GÖSA 20/OLSA variable

Note. All measurements were performed via the near field loudspeaker. During all measurements, restaurant noise was used as interfering noise at 62 dB presentation level. GÖSA = Göttingen sentence test; OLSA = Oldenburg sentence test; SRT = speech reception threshold.

Table 3. HA Condition and Set Up Parameters for the Dual-Task Speech Intelligibility Test Paradigm.

	Primary task (OLSA)	Secondary task (GÖSA)	HA condition
1	Near field	Far field	Omni + RM
2	Near field	Far field	Dir + RM
3	Far field	Near field	Omni + RM
4	Far field	Near field	Dir + RM

Note. GÖSA = Göttingen sentence test; OLSA = Oldenburg sentence test; Omni + RM = omnidirectional hearing aid microphones with remote microphone; Dir + RM = directional hearing aid microphones with remote microphone.

yielding a presentation level range from 59 to 71.3 dB SPL across all participants. The secondary task presentation level was chosen in this way to achieve speech intelligibilities clearly above 50%, but below 100% in the dual-task condition to minimize ceiling effects.

The dual-task speech intelligibility measurements for the different HA and measurement conditions were performed in balanced order (see Table 3).

After a recess, the participants performed the interleaved speech intelligibility tests. For training purposes, three adaptive measurements of SRT_{50} were performed with 2×10 sentences each in HA condition Omni + RM. The resulting minimum individual speech presentation levels of these three measurements + 2 dB were chosen for each participant as the fixed speech presentation level in the following tests, separately for near field and far field presentation. The resulting SNRs of the subsequent measurements ranged from −13.1 to −4.8 dB SNR for far field presentation and from −1.2 to 4.8 dB SNR for near field presentation. The two HA conditions Omni + RM and Dir + RM were measured in balanced order. The interleaved speech intelligibility measurements were also performed with an additional hearing condition that is not further reported in this article: front focused binaural directionality without RM.

After the completion of Visit 1, the participants used the test devices together with the RM in the second daily life period for 3 more weeks with automatic directionality of the HA microphones.

Visit 2 started again with training of the dual-task paradigm according to Table 2, except that the last measurement of the dual-task test (No. 5) was skipped.

Thereafter, all dual-task measurements from the first appointment were rerun in the same order as in the previous visit.

After performing training of the interleaved speech intelligibility test using 2×10 sentences, the same interleaved speech intelligibility measurements were carried out in the same order as before.

The presentation levels of the dual-task test and the interleaved speech intelligibility measurements were the same as in the prior visit.

The main difference between Visit 1 and Visit 2 was that the participants were acclimatized to the RM in the second visit, while they had no experience with any kind of RM in the first visit.

Statistical Analysis

The data were analyzed with a $2 \times 2 \times 2$ analysis of variance (ANOVA; factors HA condition, set up, visit) in order to reveal the main effects and direct interactions. Therefore, data distributions were checked for normality by Shapiro Wilk tests. According to Bortz (1999), a repeated measures ANOVA is robust against violations of the assumption of normality as long as the same number of observations for the dependent variables are included. Therefore, the ANOVA was still performed even if the Shapiro Wilk test indicated an abnormal data distribution.

Results

Dual-Task Speech Intelligibility Measurements

The dual-task speech intelligibility measurements were first analyzed by addressing the performance on either primary task or secondary task alone, although both tasks were performed simultaneously during the measurements. Figure 5 shows the results of the primary

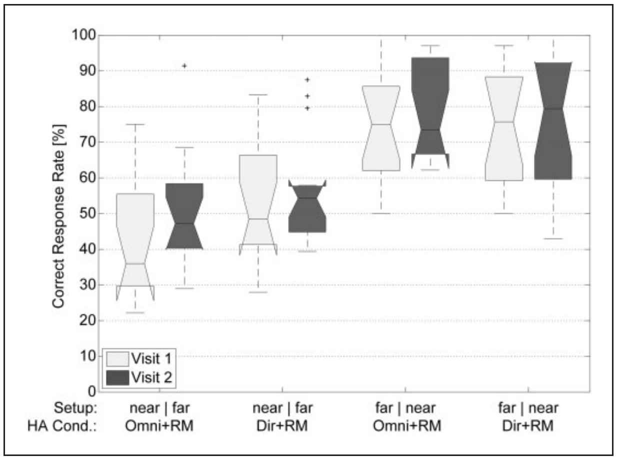

Figure 5. Boxplot of the primary task results (correct response rate of OLSA names presented as percentages) for all participants. In the figure, the setup is always described as the speaker used for the primary task, followed by the speaker used for secondary task. In the following, the setup is always denoted by the speaker position of the primary task. HA Cond. = hearing conditions; Omni + RM = omnidirectional hearing aid microphones with remote microphone; Dir + RM = directional hearing aid microphones with remote microphone.

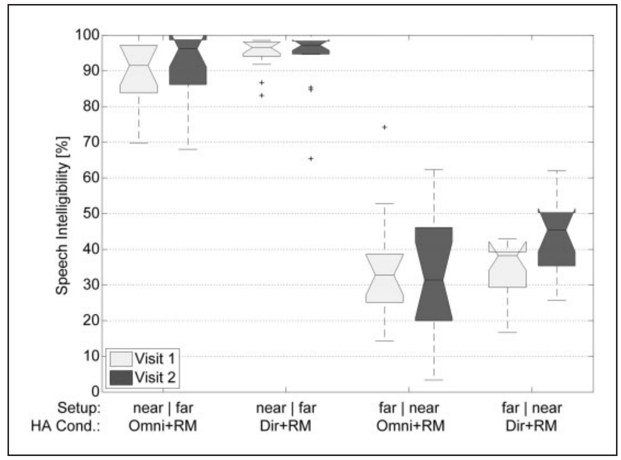

Figure 6. Boxplot of the secondary task results (speech intelligibility in percentages determined with GÖSA). Notation and comments similar to Figure 5.

task, where correct response rates of OLSA names are presented as percentages, shown as a boxplot for all participants. Results are given separately for HA condition and either near or far field presentation of the primary task. Light gray bars show the results of Visit 1, and dark gray bars for Visit 2. The HA conditions and the setup are indicated on the abscissa of the graph. Two boxes (respectively their medians) are significantly different at the 5% significance level, if their intervals (i.e., the ''notches'' of the boxes) do not overlap. The interval endpoints are indicated by the extremes of the notches, corresponding to $Median - \frac{1.57(q3-q1)}{\sqrt{n}}$ and $Median + \frac{1.57(q3-q1)}{\sqrt{n}}$ where *q1* and *q3* are the 25th and 75th percentiles of the sample data, respectively, and *n* is the number of observations. For small sample sizes, the notches might extend beyond the end of the box. Outliers are plotted separately, if they are greater than $q3 + 1.5(q3 - q1)$ or less than $q1 - 1.5(q3 - q1)$. The plotted whisker extends to the adjacent value, which is the most extreme data value that is not an outlier.

Due to Shapiro Wilk, the data of DIR + RM in Visit 2 for primary task presentation via near field and far field were not normally distributed with an error probability of 5%.

The $2 \times 2 \times 2$ ANOVA revealed a significant main effect of set up: Primary task via far field has a higher correct response rate than via near field, $F(1,14) = 63.24$, $p < .001$, and a significant main effect of visit: higher correct response rate in Visit 2 than in Visit 1, $F(1,14) = 5.62$, $p = .033$. Although there was no statistically significant main effect of HA condition, a significant interaction of Set Up $\times$ HA condition was found, $F(1,14) = 13.86$, $p = .002$: Only in near field, DIR + RM yielded higher correct response rates than in OMNI + RM.

Figure 6 shows the results of the secondary task, boxplots of speech intelligibility in percentages determined with GÖSA, averaged across participants.

Due to Shapiro Wilk, the data of DIR + RM in Visit 2 for secondary task presentation via near field and far field and the DIR + RM data in Visit 1 for secondary task presentation via near field were not normally distributed with an error probability of 5%.

The $2 \times 2 \times 2$ ANOVA revealed a significant main effect of set up: Secondary task via far field elicited a higher correct response rate than via near field, $F(1,14) = 340.85$, $p < .001$, and a significant main effect of HA condition: higher correct response rate with DIR + RM than with OMNI + RM, $F(1,14) = 7.72$, $p = .015$. There was no statistically significant main effect of visit and no factor interactions.

A central question of this study is to determine the common performance of both tasks in the dual-task speech intelligibility paradigm. The underlying idea is a capacity or resource model that shares the resources on both the primary and the secondary tasks. In such a model, the dual-task costs indicate the loss of resources that go to the respective other task: Most likely, the performance of each task drops when performing both tasks at the same time compared with the case when each task is carried out separately (Halvorson, 2013).

In the following, dual-task costs are calculated using the ''probit'' (probability units) transformation of the *differences* in correct response rates (primary task, OLSA) or speech intelligibility (secondary task, GÖSA), respectively, for each task in the dual condition

(Figures 5 and 6) compared with the single-task condition (data not shown here), according to Oberauer, Lange, and Engle (2004). The basic assumption is that a continuous variable (the "resource") is translated into percent correct by a sigmoid (e.g., logistic) function. This is done by a probit transformation, which translates the probability of a correct answer into the corresponding *z*-score of a standard normal distribution. It takes both floor and ceiling effects of the percent-correct scale into account. In this regard, it is neutral, which is especially important when not specifying a particular model for the underlying performance.

The OLSA (primary task) was measured as a single task during the training session for each subject. This measurement was used to calculate the change (decrease) in performance as a result of adding the secondary task in the dual-task measurements. We used this measurement with test devices in the focused front directionality setting of the binaural beamformer but without RM as the reference for all other dual-task measurements. For the secondary task (GÖSA), no single-task measurement was done beforehand. Instead, we assumed that an intelligibility of 100% would be achieved when performing this task separately since the presentation level during the dual-task test was 3 dB higher than the measured SRT_{50} of the last dual-task training measurement (No. 5 in Table 2). We assumed that the intelligibility at that SNR is 100% for the single task. This was not proven, but it was used as an approximation anchor value that is constant for all conditions.

We used the *difference* between the single- and dual-task conditions for calculating the *z* scores. Therefore, the outcome is "negative" when assigned to dual-task costs, that is, higher values reflect fewer dual-task costs and thus better dual-task performance. Therefore, the outcome of the transformation is labeled as *dual-task performance*.

Our assumption, in terms of the resource model, is that the total dual-task performance is the sum of both dual-task performances of the primary and secondary task, see Figure 7.

Due to Shapiro Wilk, the Visit 1 data of OMNI + RM for primary task presentation via near field (secondary task presentation via far field) and the DIR + RM data for primary task presentation via far field (secondary task presentation via near field) were not normally distributed with an error probability of 5%.

The $2 \times 2 \times 2$ ANOVA revealed significant main effects of all tested factors. HA condition: DIR + RM yielded significantly lower dual-task costs (higher performance) than OMNI + RM, $F(1,14) = 7.24$, $p = .018$. Set up: Primary task presentation via far field yielded significantly higher dual-task performance than via near field, $F(1,14) = 41.52$, $p < 0.001$. Visit: Visit 2 yielded significantly higher dual-task performance than Visit 1, $F(1,14) = 7.25$, $p = .017$. No statistically significant interactions of factors could be found.

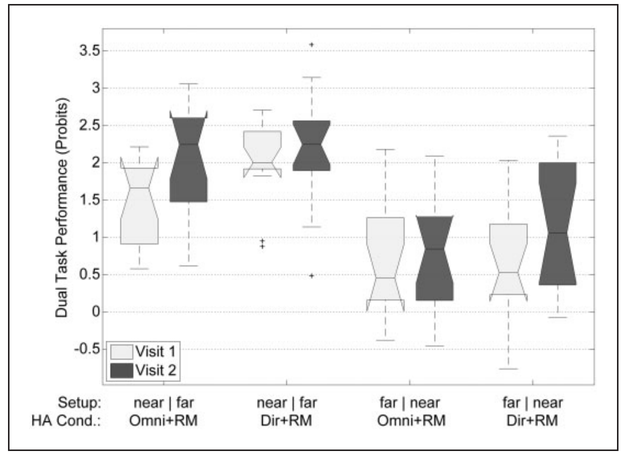

Figure 7. Boxplot of the dual-task performance. Higher values reflect less dual costs, thus better performance. Notation and comments similar to Figure 5.

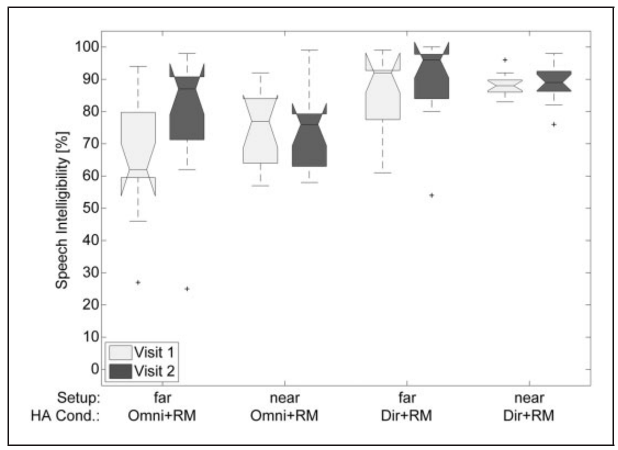

Figure 8. Boxplots of the interleaved measurements results. Notation and comments similar to Figure 5.

Interleaved Speech Intelligibility Measurements

The interleaved speech intelligibility measurements were performed with individually chosen fixed SNRs. These SNRs were determined in a premeasurement phase with Omni + RM HA condition. Figure 8 shows the results of the interleaved speech intelligibility measurements of OLSA separately for presentation in near or in far field. Results from Visit 1 are given in light gray; results from Visit 2 are given in dark gray.

Due to Shapiro Wilk, the data of DIR + RM in Visit 1 and 2 for speech presentation via far field and the OMNI + RM data in Visit 2 for speech presentation via far field were not normally distributed with an error probability of 5%.

The $2 \times 2 \times 2$ ANOVA revealed a significant main effect of HA condition: higher speech intelligibility with DIR + RM than with OMNI + RM, $F(1,14) = 45.00$, $p < .001$, and a significant main effect of visit: higher speech intelligibility in Visit 2 than in Visit 1, $F(1,14) = 6.32$, $p = .025$). In addition, a statistically significant interaction of factors set up and visit could be found, $F(1,14) = 7.88$, $p = .014$: no speech intelligibility difference between visits in near field but better speech intelligibility in Visit 2 compared with Visit 1 in far field.

Discussion

The current study investigated the speech intelligibility benefit of directional HA microphones in the context of combined usage of HAs and RMs. Communication situations requiring either shared attention (dual-task speech intelligibility test) or shifting attention (interleaved speech intelligibility test) were simulated in the laboratory measurements.

The shared attention communication situation (dual-task speech intelligibility) was first analyzed by a separate analysis of the primary and secondary task data. This analysis of the primary task results showed that the fixed presentation levels of 65 dB SPL for near field presentation and 70 dB SPL for far field presentation of the primary task material (OLSA sentences) were well chosen to avoid any floor or ceiling effects in the primary task. The significantly higher primary task correct response rate for presentation via far field compared with presentation via near field can be explained by the higher speech presentation levels and an additional amplification provided by the RM setting: The RM automatically amplifies the receiver gain by minimum 10 dB up to 30 dB when the environmental noise exceeds a threshold between 50 and 60 dB SPL (Johnson & Seaton, 2011; Thibodeau, 2014). The main effect of visit may have been caused by both further training of the task for the second visit and an effect of RM acclimatization. Since no test–retest data are available without intermediate RM acclimatization, these aspects cannot be separated. The results of the additional HA condition without RM were not reported and further analyzed in this article. However, they can be used to motivate that the main effect of visit seems to be caused by RM acclimatization more than by training of the task: The results without RM usage were similar across visits one and two. Since the conditions were measured in random order, there is no reason that a training effect should be visible only in some of the measured conditions and therefore acclimatization to RM seems to cause the visit differences. Maybe this is due to acclimatization to the higher gain that is applied to the signals when the RM is used in these clinical settings (10 dB amplification relative to the HA microphone output). The statistically significant interaction of factors HA condition and set up shows that the positive effect of HA microphone directionality was only present for near field speech presentation of the primary task.

The speech presentation levels of the secondary task were individually chosen but fixed across measurement conditions. The separate analysis of secondary task results showed that in HA condition, DIR + RM ceiling effects may have occurred for far field presentation of the secondary task and also in HA condition OMNI + RM during the second visit. This may have been the reason that no significant visit effect could be found in the secondary task. The significant main effect of set up on secondary task performance may also be due to the additional amplification of the RM signal compared with HA microphone signals like in the primary task. The main effect of HA condition showed the benefit from the directional microphone settings both in near field and in far field secondary task presentation. The benefit in near field presentation was smaller compared with the benefit from directional HA microphones found by Jones and Rakita (2016), since speech intelligibility was not measured in a dual-task paradigm in that study (10% instead of 26% average benefit). The benefit in far field presentation may be caused by the noise suppression effect of the directional microphone HA condition. This effect could only be seen for the secondary task and in the dual-task cost analysis since the measurement accuracy of the primary task was smaller. The number of words tested in one measurement was approximately 34 in the primary task (average number of presented OLSA sentences and respective target names) and approximately 100 in the secondary task (20 GÖSA sentences with 3–7/average five words each).

To investigate the shared attention/speech intelligibility aspects of the dual-task speech intelligibility measurements, the results were analyzed with regard to common performance of primary and secondary tasks. The statistically significant main effect of HA condition indicates that the directional microphone settings of the HAs enables better speech intelligibility regardless of whether the talker is near or far. It shows that the directional microphone not only improves speech perception in the near field but also for a distant speaker transmitted to the HA via RM. In this case, the directional microphone acts as an additional means of noise suppression. The main effect of set up indicates that the different levels of task difficulties are influenced by the distance from which the tasks are presented. When transmitting the secondary task speech material (GÖSA sentences) from far field via RM, the performance was much better than when it was received from near field with the microphones of the HA. This was due to the default mixing factor at the input stage of the HA that was set to 10 dB amplification of the RM signal versus the HA microphone signal. The result also shows that the secondary task seems to be the determining factor in the common performance of the dual-task speech intelligibility test. Thus, the positive main effect of HA condition DIR + RM in the common dual-task performance is due to both the directionality in near field and noise suppression in far field, similar to the secondary task performance. The secondary task performance in Visit 2 was higher than in Visit 1, although the difference was not significant. Considering the common performance of primary and secondary tasks, the visit differences were statistically significant similar to those for primary task performance due to additional training of the measurement procedures and RM acclimatization.

The missing interaction effect for both HA condition and set up supports the additional benefit provided by the directional HA microphone setting regardless of where both tasks are presented from. DIR + RM was shown to be beneficial in both far and near field regarding common performance.

The results of the shifting attention communication situation (interleaved speech intelligibility) show that the individually chosen speech presentation levels were adequately chosen avoiding floor and ceiling effects except for the DIR + RM condition. The intelligibilities of near field and far field presentation did not differ (factor set up) since the presentation levels were chosen separately for near field and far field presentation based on pretests resulting in 50% intelligibility in the interleaved speech intelligibility task. The average presentation level difference between near and far field presentation was −11.2 dB. Similar to the dual-task results, the main effect of HA condition showed that the benefit from HA microphone directionality also affects far field speech intelligibility due to noise suppression. The benefit of directional HA microphones compared with omnidirectional in near field speech intelligibility (12% in average) is of similar magnitude like that found in the secondary speech intelligibility task in the dual-task paradigm and therefore less than found by Jones and Rakita (2016). Similar to the dual-task speech intelligibility measurements, the better performance of Visit 2 compared with Visit 1 may be due to both additional training of the task before Visit 2 and acclimatization to RM. It seems that acclimatization to RM is the prominent influence since the performance differences between Visit 2 and visit 1 are obviously larger for far field performance than for near field performance and there is also a statistically significant interaction effect of factors and visit although there is some ceiling for far field performance of HA condition DIR + RM.

Conclusions

The study investigated speech intelligibility with severely impaired listeners in complex listening scenarios where either shared or shifting attention was needed when using HAs together with a RM. It could be shown that (a) directional HA microphones yielded higher speech intelligibility than omnidirectional microphone settings both in near and in far field if highly accurate speech intelligibility measures are used in a dual-task or interleaved intelligibility measurement. This is due to the directionality effect in the near field and the noise suppression effect in the far field. (b) The benefit from directional HA microphones is higher in the interleaved measurements requiring shifting attention compared with the dual task asking for shared attention. (c) There is some evidence that speech intelligibility in these complex listening environments with HAs and RM increases with acclimatization to RM usage.

Acknowledgments

The authors thank Maria Klatte for pointing us to advanced methods for evaluating dual-task measurements. Parts of this research were presented at the 6th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, August 23–25, 2017.

Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Matthias Latzel and Hans E. Mülder are employees of Sonova AG.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by Sonova AG, Stäfa, Switzerland.

ORCID iD

Kirsten C.Wagener http://orcid.org/0000-0003-2858-2789

References

American Academy of Audiology Clinical Practice Guidelines. (2011). *Remote microphone hearing assistance technologies for children and youth from birth to 21 years*. Retrieved from http://www.audiology.org

American National Standard. (2010). *Method of measurement of performance characteristics of hearing aids under simulated real-ear working conditions* (ANSI/ASA S3.35:2010). Melville, NY: Acoustical Society of America.

Anderson, K. L., & Goldstein, H. (2004). Speech perception benefits of FM and infrared devices to children with hearing aids in a typical classroom. *Language, Speech, and Hearing Services in Schools*, *35*, 169–184.

Bertrand, A., & Moonen, M. (2009). Robust distributed noise reduction in hearing aids with external acoustic sensor nodes. *EURASIP Journal on Advances in Signal Processing*, *2009*, article ID 530435, 14 pages.

Blazer, C. (2007). Improving the classroom environment: Classroom audio technology (Miami-Dade County Public Schools Information Capsule, Research Services, 607).

Miami, FL: Miami-Dade County Public Schools Information.

Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5th ed.). Berlin, Germany: Springer.

Brand, T., & Kollmeier, B. (2002). Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *Journal of Acoustical Society of America*, *111*, 2801–2810.

Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acoustica*, *86*, 117–128.

De Ceulaer, G., Bestel, J., Mülder, H. E., Goldbeck, F., De Varebeke, S. P. J., & Govaerts, P. J. (2016). Speech understanding in noise with Roger Pen, Naida CI Q70 processor, and integrated Roger 17 receiver in a multi-talker network. *European Archives of Oto-Rhino-Laryngology*, *273*, 1107–1114.

Dillon, H. (2012). *Hearing aids* (2nd ed.). Sydney, Australia: Boomerang Press.

Gößling, N., Marquardt, D., & Doclo, S. (2017). Performance analysis of the extended binaural MVDR beamformer with partial noise estimation in a homogeneous noise field. In *Proceedings of the joint workshop on hands-free speech communication and microphone arrays (HSCMA)* (pp. 1–5). San Francisco, CA: IEEE Signal Processing Society, IEEE.

Halvorson, K. M. (2013). *What causes dual-task costs?* (Doctoral thesis, University of Iowa). Retrieved from http://ir.uiowa.edu/etd/1331

International Electrotechnical Commission. (2011). Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index (IEC 60268-16). Geneva, Switzerland: International Electrotechnical Commission.

Johnson, C. D., & Seaton, J. B. (2011). *Educational audiology handbook*. New York, NY: Delmar Cengage Learning.

Jones, C., & Rakita, L. (2016). A powerful noise-fighting duo: Roger and Phonak directionality. *Phonak Field Study News*. Retrieved from http://www.phonakpro.com/com/b2b/en/evidence.html

Killion, M. C. (2004). Myths about hearing in noise and directional microphones. *The Hearing Review*, *11*(2). Retrieved from https://www.etymotic.com/media/publications/erl-0051-2004.pdf.

Killion, M. C., & Niquette, P. A. (2000). What can the pure-tone audiogram tell us about a patient's SNR loss? *The Hearing Journal*, *53*(3), 46–53.

Kim, J. S., & Kim, C. H. (2014). A review of assistive listening device and digital wireless technology for hearing instruments. *Korean Journal of Audiology*, *18*(3), 105–111.

Kochkin, S. (1996). Consumer satisfaction & subjective benefit with high performance hearing aids. *The Hearing Review*, *3*(12), 16–26.

Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M. A., Uslar, V., Brand, T., & Wagener, K. C. (2015). The multilingual matrix test: Principles, applications, and comparison across languages: A review. *International Journal of Audiology*, *54*(S2), 3–16.

Kollmeier, B., & Wesselkamp, M. (1997). Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *Journal of Acoustical Society of America*, *102*(4), 2412–2421.

Oberauer, K., Lange, E., & Engle, R. W. (2004). Working memory capacity and resistance to interference. *Journal of Memory and Language*, *51*, 80–96.

Plomp, R. (1977). Acoustical aspects of cocktail parties. *Acustica*, *38*, 186–191.

Szurley, S., Bertrand, A., Van Dijk, B., & Moonen, M. (2016). Binaural noise cue preservation in a binaural noise reduction system with a remote microphone signal. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, *24*(5), 952–966.

Thibodeau, L. (2014). Comparison of speech recognition with adaptive digital and FM wireless technology by listeners who use hearing aids. *American Journal of Audiology*, *23*, 201–210.

Vroegop, J. L., Dingemanse, J. G., Homans, N. C., & Goedegebure, A. (2017). Evaluation of a wireless remote microphone in bimodal cochlear implant recipients. *International Journal of Audiology*, *56*, 643–649.

Wagener, K., Brand, T., Kühnel, V., & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I-III: Design, Optimierung und Evaluation des Oldenburger Satztests [Development and evaluation of a German sentence intelligibility test I-III: Design, optimization, and evaluation of the Oldenburg sentence test]. *Zeitschrift für Audiologie*, *38*(1–3), 4–15, 44–56, 86–95.

Wagener, K. C., & Brand, T. (2005). Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters. *International Journal of Audiology*, *44*(3), 144–157.

Wagener, K. C., Hochmuth, S., Ahrlich, M., Zokoll, M. A., & Kollmeier, B. (2014). Der weibliche Oldenburger Satztest [The female Oldenburg sentence test]. In *Proceedings of 17 Jahrestagung der Deutschen Gesellschaft für Audiologie*, CD-ROM, 4 *pp*, Oldenburg, Germany. Retrieved from www.dga-ev.com.

Wilson, R. H., McArdle, R. A., & Smith, S. L. (2007). An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *Journal of Speech Language and Hearing Research*, *50*(4), 844–856.

Wolfe, J., Duke, M. M., Schafer, E., Jones, C., Mülder, H. E., John, A., & Hudson, M. (2015). Evaluation of performance with an adaptive digital remote microphone system and a digital remote microphone audio-streaming accessory system. *American Journal of Audiology*, *24*(3), 440–450.

Wolfe, J., Morais, M., Schafer, E., Mills, E., Mülder, H. E., Goldbeck, F., ... Lianos, L. (2013). Evaluation of speech recognition of cochlear implant recipients using a personal digital adaptive radio frequency system. *Journal of American Academy of Audiology*, *24*(8), 714–724.

Wolfe, J., Schafer, E. C., Heldner, B., Müdler, H., Ward, E., & Vincent, B. (2009). Evaluation of speech recognition in noise with cochlear implants and dynamic FM. *American Journal of Audiology*, *20*(7), 409–421.