



# The effect of symmetric and asymmetric directional binaural listening on speech understanding with surrounding cocktail party noise

Luca Giuliani<sup>1,2</sup> · Luca Brayda<sup>2</sup>

Received: 18 September 2018 / Accepted: 20 March 2019 / Published online: 26 March 2019  
© The Author(s) 2019

## Abstract

The development of new augmenting hearing devices is reducing the separation between smart listening technologies and hearing aids. They both alter the way people perceive surrounding acoustic sources. One of their common goals is reducing the effect of noisy or crowded environments on listening capabilities. However, it is still unclear how much spatial processing, i.e. conditioning sounds depending on their direction of arrival, can be beneficial to binaural feedback. This study investigates the effect of binaural symmetrical and asymmetrical spatial processing on speech comprehension in a noisy environment, when the listeners are people with no hearing impairment. 15 participants sat at the center of four speakers positioned at 0°, 90°, 180° and 270°, listening and repeating full sentences coming from the speakers, while a competing cocktail party noise was reproduced. The task was repeated in four listening conditions: Free ear, Omnidirectional, Directional and Asymmetric, which differed by the presence and kind of spatial processing performed by a pair of glasses equipped with microphone arrays. We found that Directional and Asymmetric fittings performed similarly for frontal sources, but were significantly more effective than Free ear and Omnidirectional. Then, Asymmetric condition showed to be better than Directional for speech coming from one of the sides, but worse than Omnidirectional and Free ear. Overall, asymmetrical fitting could be exploited by augmenting hearing devices to ease communication with multiple talkers in noisy environments or to exclude or reduce the impact of unwanted noise coming from specific directions.

**Keywords** Hearing · Augmented hearing · Microphone arrays · Hearing aids · Binaural cues

## 1 Introduction

Partial hearing loss is a common condition: 360 Million people in the world suffer from disabling hearing loss (WHO 2016). Recent technological advancements brought on the market hearing aids that improve speech understanding during conversations with a single human talker, or while watching television or during leisure activities (European Hearing Instrument Manufacturers Association 2015, 2016). Nonetheless, very common acoustic scenarios such as noisy or crowded environments, or the presence of multiple surrounding talkers around the listener, are still problematic (McCormack and Fortnum 2013).

Typically, the brain extracts features from sounds (called binaural cues) and uses them to spatially segregate different sound sources or to estimate their directions of arrival, therefore solving the so-called *cocktail party problem* (Cherry 1953). The outcome is an attentional focus of the listener on a single voice or sound in a complex acoustic scenario.

Partial hearing loss often leads to a deterioration in the capability of using such cues and locating sounds in space (Brimijoin et al. 2010). For this reason, modern hearing aids adopt directional filtering to reduce the complexity of the acoustic scenario and increase speech understanding. Apart from noise reduction and speech denoising and de-reverberation algorithms (Chung 2004; Klasen et al. 2007; Luts et al. 2010), the main feature of modern hearing aids is the ability to use two or more microphones to filter sounds in different ways on the basis of their direction of arrival, using beamforming techniques (Doclo et al. 2010; Doclo and Moonen 2003). The use of microphone arrays offers great flexibility, since it allows to obtain spatial filters with different segregation power and steering direction without

---

✉ Luca Giuliani  
luca.giuliani@iit.it

<sup>1</sup> DIBRIS - University of Genoa, Genoa, Italy

<sup>2</sup> RBCS - Fondazione Istituto Italiano di Tecnologia, Genoa, Italy

hardware modifications, enhancing the Signal to Noise Ratio (SNR) in the direction of interest.

Previous studies (Doclo et al. 2009) showed that binaural directional sound filtering can increase the comprehension of frontal speech (i.e. when the source of interest is in front of the person) in noise for people with and without hearing impairment. Such solution, however, assumes that the same kind of spatial filter is applied to both ears, where filtering can be defined ‘symmetric’. The question is whether increasing the spatial selectivity of hearing aids (i.e. pointing as much as possible a single source, which translates in narrow main lobes of the beam pattern, see (Brandstein 2000)), is always beneficial, or, instead, if there are situations in which the position of target sound sources could negatively affect speech comprehension. In fact, one can also design ‘asymmetric’ configurations, defined as the combination of omnidirectional listening on one ear and directional listening on the other ear (see Fig. 1).

The rationale is that it is entirely possible that asymmetric fitting could be effective in excluding lateral sound sources to the same side of the directional ear, rather than in focusing on lateral sound sources to the same side of the omnidirectional ear, with minimal impact on listening capabilities for the location of other sound sources.

Most studies investigated the influence of asymmetrical fitting on speech understanding of frontal sound sources only, with hearing impaired people (Devocht et al. 2016; Hornsby and Ricketts 2007; Kim and Bryan 2011; Mens 2011; Ricketts and Picou 2013).

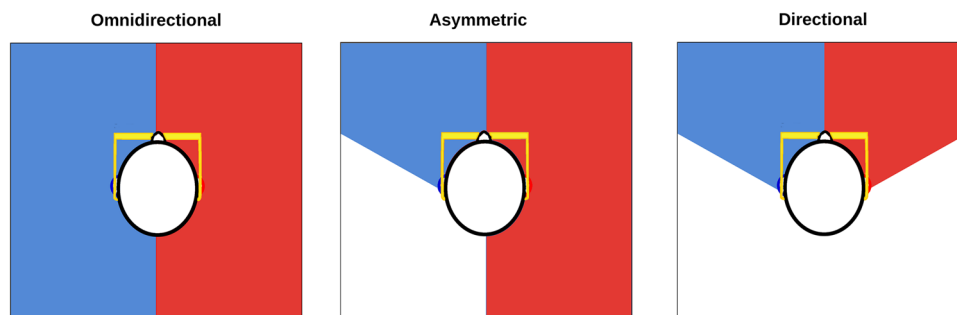
Asymmetrical fitting appears to be more efficient than omnidirectional fitting, but contradictory findings emerge on the comparison between asymmetric and purely directional (i.e. symmetrical) fitting. In fact some authors (Devocht et al. 2016; Hornsby and Ricketts 2007; Ricketts and Picou 2013) found that symmetric directional fitting significantly improves speech comprehension of frontal sound sources as compared to asymmetric, while others (Kim and Bryan

2011; Mens 2011; Picinali and Prosser 2010) could find only a slight, non-significant, difference between the two configurations.

The common way used to assess the effectiveness of asymmetric configurations so far consists in testing speech reception capabilities of participants when the target speech is always located in front of the person and the competing noise is coming from different directions: only a few studies (Hornsby and Ricketts 2007; Picinali and Prosser 2010) investigated the effectiveness of asymmetric fitting for lateral sources. Non-frontal sound sources could be an important source of information in evaluating directional filtering solutions. For people with hearing aids it is very frequent to participate in conversations with more than one speaker or in which the speaker is not in front.

Some algorithms exist that automatically adapt to the acoustic scenario, for example estimating the position of the speaker and updating the steering direction of a hearing aid, or using Blind Source Separation (Adiloğlu et al. 2015; Kamkar-Parsi et al. 2014; Luts et al. 2010). However, it has been shown that in complex scenarios these solutions can fail, with a negative impact on the speech comprehension of the person with hearing impairment (Luts et al. 2010).

Another aspect concerns the tested population. Truly, most studies test people with hearing impairments, since they receive most of the advantages derived from the development of new generations of hearing aids. However, with the recent advent of augmented hearing technologies, it is interesting to study how directional listening (both symmetrical and asymmetrical) affects normally hearing people or at least people with a moderate degree of hearing impairment. There are a number of use cases where directional listening may be beneficial to persons without chronic damage to the hearing system. These include those for which speech understanding is seriously challenged: crowded places (open-air concerts, conversations along roads full of traffic, restaurants at meal time) or very noisy situations (workplaces close to



**Fig. 1** Schematic representation of omnidirectional configuration (left), directional configuration (right) and asymmetric configuration (center). In the omnidirectional condition the ears perceive all sounds coming from every direction, while in the directional configuration

the intensity of sounds coming from the sides and the back are heavily attenuated, while the frontal sounds are preserved. The asymmetric configuration is a combination of the two previous techniques. Picture inspired by Devocht et al. (2016)

machineries or aircrafts). Such scenarios can represent, at the sensory level, *temporary disabilities* because the signal-to-noise ratio is very low. Solutions that improve speech intelligibility, properly scaled to a population that has little or no permanent impairment, can be similar to those adopted with hearing aids. Testing psycho-acoustical properties of novel technologies on normally hearing participants, therefore, allows to decouple the effect of a certain solution on the kind of hearing impairment. It can therefore have a meaning per se (Ricketts and Picou 2013).

In fact, Picinali and Prosser (2010) investigated the impact of asymmetric microphone configuration on normally hearing participants in a simulated environment for frontal and lateral sources, with the Ambisonics technique (Gerzon 1973, 1974).

Their findings about lateral sources are quite interesting: it seems that the asymmetric configuration was significantly more effective than both omnidirectional and symmetric directional for sounds coming from the side. Yet, that study has been conducted in a simulated environment, with microphones modeled as ideal cardioids and microphone positioning coincident in a single point in space, i.e. without considering interaural distance.

Here we study the impact of four symmetric and asymmetric binaural configurations on normally hearing participants in an ecological environment. We verify how speech reception in noisy conditions (cocktail party noise) is affected by these configurations for frontal, lateral and posterior speech sound sources. We contribute to solve the debate about whether or not asymmetric fitting can increase understanding of lateral speech sources without decreasing that of frontal speech sources.

Our hypothesis is that the asymmetric fitting can enhance the speech reception for lateral and posterior stimuli, without significantly affecting the comprehension of frontal sources.

We also want to verify if the asymmetric configuration can exceed the performance of symmetric omnidirectional for lateral speech, like found by Picinali and Prosser. Our hypothesis is that the asymmetric configuration is at least comparable to the omnidirectional configuration for lateral speech coming from the side of the omnidirectional listening ear.

## 2 Methods

The purpose of this study is to investigate speech reception with normally hearing participants in function of different directional filtering techniques, in an environment where cocktail party affects a desired target speech input, where both target speech and noise come from multiple directions. We tested different filtering techniques by means of beamforming algorithms applied to wearable microphone arrays.

### 2.1 Beamforming algorithm on a wearable embedded system

The extent to which beamforming techniques are effective greatly depends both on the array geometry and dimensions and on the beamforming algorithm. Concerning geometry, it is well known that linear equispaced arrays are as much directive as the number of microphone increases. If the sensors are linearly spaced, the array is as much directional at low frequencies as the first and last microphone are farther apart. Similarly, the array is as much directional at high frequencies as two adjacent microphones are closer (Brandstein 2000).

This can be a drawback if technologies, such as hearing aids, tend to be smaller and smaller, that is with an inevitably low number of microphones (Chung 2004). The problem is partially compensated by a more desired directivity at high frequencies, i.e. where the clinical consequences of hearing loss are more prominent.

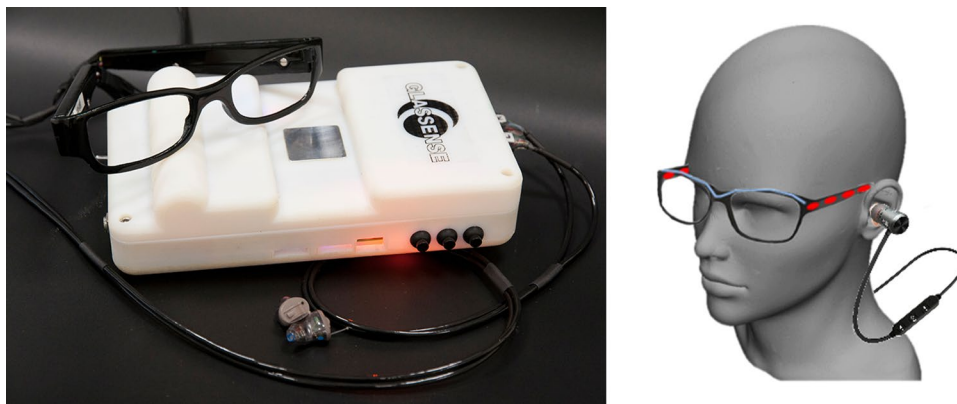
The device used in this study, named Glassense (Brayda et al. 2015), exploits two arrays of microphones to spatially filter sounds in the frequency range of human voice. That is, the array dimensions, the number and spacing of the microphones are compatible with a desired directivity in the range of 250–4000 Hz, which are the frequencies at which the human voice content is more dense (Killion and Mueller 2010). Moreover, this frequency range overlaps with that used in studies with hearing impaired persons, who generally adopt hearing aids with a bandwidth limited to 8 KHz, as done in Hornsby and Ricketts (2007) and Mens (2011).

Glassense has the shape of a pair of glasses, with the microphones arrays fixed to the temples and connected to a portable elaboration board. The device allows the listener to use the head motion as a spatial selector, indicating the direction of the desired listening focus. In other words, the directional focus is fixed to perform best (i.e. to be more spatially selective) when the person ‘looks’ at the target speech source, similarly to what is depicted in Fig. 1 (right). Figure 2 shows the attenuation applied by our directional filter with respect to the direction of arrival of the sounds at various frequencies.

The acoustic filtering is based on a bilateral superdirective beamforming technique, which exploits two arrays of  $L=0.1$  m length with  $N=4$  omnidirectional MEMS microphones each, sampled at 16 KHz. The frequency band at which the array is more directional ranges from 250 Hz to 4 kHz. This ensures that the band is well below the frequency at which spatial aliasing occurs ( $f_{alias} = 5.2$  kHz), obtained by applying the known formula  $f_{alias} = \frac{c \cdot (N-1)}{2 \cdot L}$ , where  $c$  is the speed of sound. The directivity index (DI) has a mean value of 8.6 dB over such frequency band.

Once the sound is captured by the microphone arrays, it is sent to a portable elaboration board. The board consists in

**Fig. 2** Glassense system. The audio captured by the microphones is sent to the elaboration board, which performs filtering and sends the resulting signal to two monaural earphones. Each 4-elements microphone array sends the beamformed audio to the ipsilateral, closest earphone



a MIYR Z-turn, a low-cost Linux-based development board based on ARM processing system. The sound recorded by the device is filtered by the board and sent to the user by means of earphones in less than 20 ms.

## 2.2 Participants

We recruited 15 people with normal hearing (mean age 26, range 22–36 year). Each of them passed an audiometric test (Inventis Piccolo Plus), where the minimal pure tone individual threshold between 500 Hz and 4 KHz (i.e. the frequency range in which the microphone array is directive) had to be above 20 dBHL.

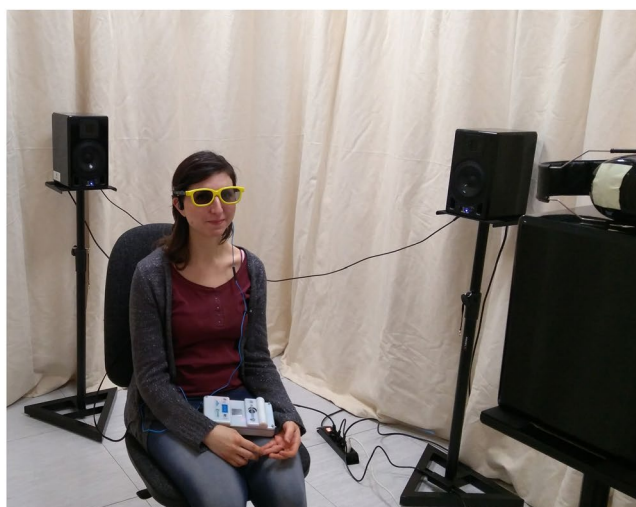
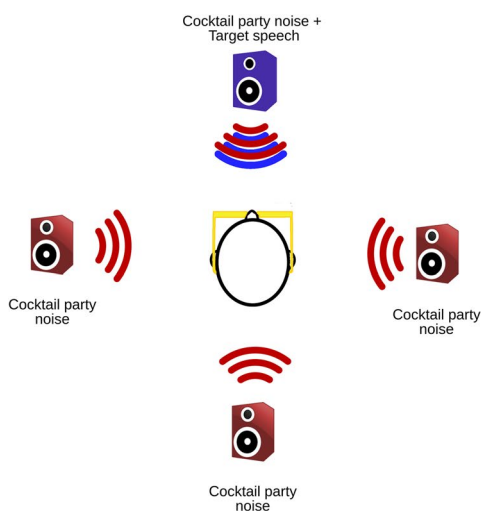
## 2.3 Setup

The setup consisted in 4 Adam A5 speakers placed at 1 m distance from the participant's head center (see Fig. 3): one

on the front ( $0^\circ$ ), one on the back ( $180^\circ$ ) and two on both sides ( $90^\circ$ – $270^\circ$ ). The speakers were connected to a laptop by means of an Asus Xonar U7 soundboard. A Python script controlled the sound reproduced by the speakers and in particular the intensity of the background noise and of the target sentences. At the beginning and at the end of each experimental session the sound level of the speakers were checked by a Delta Ohm HD2010 sound meter.

The whole experiment has been executed inside a  $3.5 \times 2.5$  m area surrounded by HOFA STUDIO Acoustic Curtains (absorption coefficient  $> 0.8$  for frequencies higher than 500 Hz) within a larger room of size  $5 \times 8.8$  m. The curtains allowed to separately analyze the contribution of each of the four speakers, since they minimized the contribution of reverberation.

We used the Glassense device to filter in real time the four sound sources and send the processed sound signals to both ears by means of Philips SHE3590PP earphones. The two



**Fig. 3** A schema of the setup used to estimate speech reception in noise (left) and a participant sitting on a fixed chair at the center of the square, facing the frontal speaker (right). Acoustic curtains sur-

rounded the setup, to isolate the participant from the external environment and reduce the effect of sound reflection

arrays in the temples could be programmed independently to function in two modes: either the array was directional (all four microphones of a single temple were used) or it was omnidirectional (only the microphone closest to the ears was used), see Figs. 4 and 5.

In the conditions in which it was requested, the participants wore the Glassense, listening to the surrounding sounds by the device earphones with the ears covered by noise isolating headphones (3M PELTOR Optime 105 Ear-muffs H10A) to avoid external sounds leakage.

## 2.4 Stimuli

The target speech stimuli consisted of 17 lists of 20 Italian sentences. We decided to avoid single words as stimuli in estimating speech reception thresholds in cocktail party noise, but rather to choose short sentences, which are more similar to common listening situations (Canzi et al. 2016; Devocht et al. 2016). Since the designed protocol, with four conditions, required a large set of stimuli (340 sentences), we joined sentences with similar content and structure from databases used by other authors. 6 out of 17 lists were composed using sentences presented in Cutugno et al. (2005) and ten lists came from (Bocca and Pellegrini 1950), from which we excluded some specific sentences containing old and disused terms. An audio file is available in the additional material. To reach the right quantity of stimuli, we created one more list imitating the available material. All the sentences used in the study were composed by 4–6 common words. All the stimuli have been recorded by a unique male speaker and have been equalized to obtain a SNR against the cocktail party noise of 0 dB.

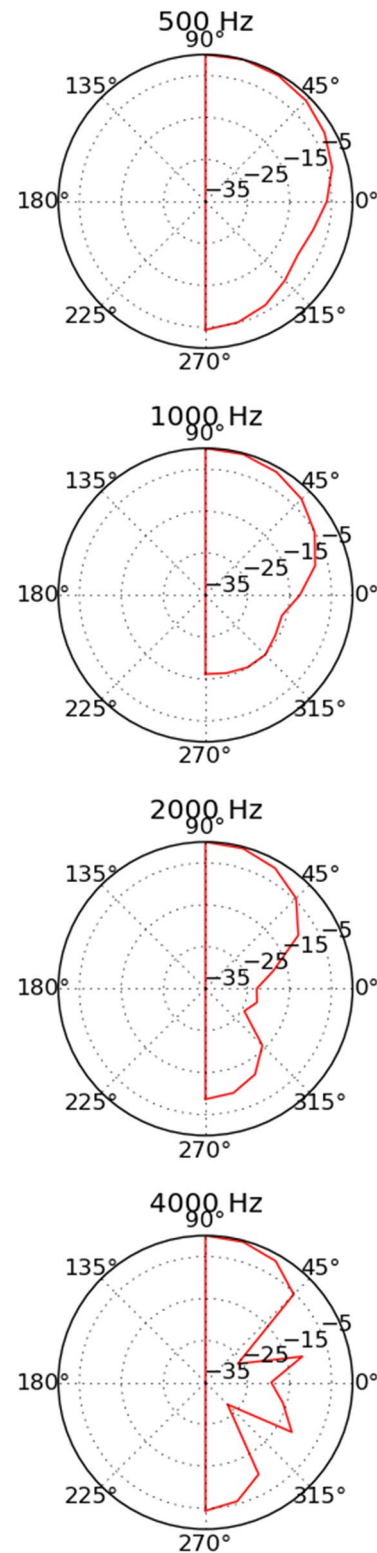
The cocktail party noise, that competed with the target stimuli, was a recording of a real cocktail party environment, reproduced by all four speakers, at a fixed level of 65 dBA. An audio file is available in the additional material. Such sound intensity was measured by placing the phonometer where the head of the participants was located. Importantly the cocktail party noise was ecologically played also by the speaker that contained the target stimulus, ensuring that noise was coming from four directions in all conditions.

All the stimuli used in the study had a bandwidth of 20 KHz.

## 2.5 Procedure

The participants comfortably sat at the center of a square, where four speakers (see Fig. 3) were placed at the corners. The participant had to face the frontal speaker, to listen to the sentences and to repeat them while they arrived from different locations, at different levels of SNR where the noise was cocktail party.

BPP single array in Free Field



**Fig. 4** Polar plot of Glassense directional filtering beam power pattern at various frequencies (90° = Front, -270° = Back)



**Fig. 5** A participant listening through the microphone arrays embedded in the Glassense device with noise isolating headphones, to avoid external sounds leakage

Each participant underwent four independent audiometries, one for each of the four locations of the target speech. Even so, the stimuli were randomized across speaker location. This minimized possible biases due to attentional focus on known directions of arrival.

The independent variable of the study consisted in the listening modality, that defined four listening conditions: (1) Free ear: listening with no device or add-on (2) Omnidirectional: listening through the Glassense, set in omnidirectional mode, that is binaural listening through one omnidirectional microphone per temple. The microphones were those above the pinna: the microphone on the left/right temple inputs to the left/right ear through a monaural earphone (see Fig. 4). The signal being spectrally filtered between the frequency range of the device (250 Hz–4KHz), with no directional processing. (3) *Directional* listening through the Glassense set in directional mode, that is binaural listening through four omnidirectional microphones per temple, the input of which is combined through superdirective beamforming, then spectrally filtered between the frequency range of the device. The beamformed output from the left/right temple inputs to the left/right ear through monaural earplugs. (4) *Asymmetric* listening through the Glassense set in directional mode on the left temple and in omnidirectional

mode on the right temple. This is equivalent to putting the left ear in *Directional* condition and the right ear in *Omnidirectional* condition.

The dependent variable was the Speech Reception Threshold (SRT), defined as the value of the Signal to Noise Ratio at which the participant can correctly guess 50% of the presented stimuli.

Each condition has been evaluated with four 20-trial iterative roving level audiometries (one for each direction). Each staircase procedure allowed to estimate a Speech Reception Threshold. We obtained for each participant 16 estimated SRT values (4 conditions  $\times$  4 directions).

More specifically, the staircase method was similar to the one presented in Canzi et al. (2016). The first of the 20 trials, for each direction, consisted in a sentence at the same intensity as the competing noise (65 dB SPL), i.e. at an SNR of 0 dB. The SNR of every stimulus after the first was calculated on the basis of the previous guess (on the same direction). Correct answers led to lower SNR on the next stimuli and vice versa, in order to converge to the 50% threshold. In a pilot study we verified that 20 trials were sufficient to converge on a reliable estimate of the SRT: the estimate of the SNR value corresponding to the SRT was obtained by averaging the SNRs of the cocktail party noise of the last 7 trials. The intensity of the competing noise was fixed during the whole experiment: instead, we varied the target stimulus sound intensity, as also done in Canzi et al. (2016).

The sentences arriving from the four directions were randomly alternated (i.e. the participant was not aware where the next sentence would come from, nor was able to predict that).

## 2.6 Data analysis

We analyzed the distributions of SRT values: each distribution pooled the estimated SRT values of all participants for each of the four conditions and each of the four directions. The goal of the study is to seek whether different conditions or directions affected such distributions. All SRT exhibited normal distributions (Shapiro–Wilk test). The SRT means and s.d. are reported in Table 1.

A first two-way repeated measures ANOVA showed that SRTs were affected by both listening condition and direction of arrival of the stimulus ( $p$  value  $\ll 0.001$ ). Four one-way repeated measures ANOVA, one for each direction of arrival, have been performed to test for significance across listening conditions ( $p$  values  $\ll 0.001$ ), followed by paired  $t$  tests within each direction. Since we made multiple comparisons, all  $t$  tests were corrected with False Discovery Rate for each stimuli direction. The statistical significances are listed in Tables 2, 3, 4 and 5.

In the following, we assume that two conditions are defined as different only if they correspond to  $p$

**Table 1** Mean SRT values and standard deviations between listening conditions and directions of arrival of the stimuli

	Front (SD)	Right (SD)	Back (SD)	Left (SD)	Mean
Free ear	− 4.25 (2.54)	− 9.56 (1.91)	− 3.61 (1.47)	− 9.84 (1.45)	− 6.81
Omnidirectional	− 4.73 (1.83)	− 8.10 (1.83)	− 5.71 (1.19)	− 8.02 (2.03)	− 6.64
Directional	− 6.28 (1.68)	1.34 (2.10)	0.90 (1.64)	0.99 (1.99)	− 0.76
Asymmetric	− 6.73 (1.91)	− 6.37 (1.53)	− 3.77 (1.30)	0.24 (1.11)	− 4.16
Mean	− 5.50	− 5.67	− 3.05	− 4.16	

**Table 2** *p* values of SRT for frontal stimuli

Front	Free ear	Omnidirectional	Directional	Asymmetric
Free ear	–	NS	<0.05*	<0.05*
Omnidirectional		–	<0.05*	<0.05*
Directional			–	NS
Asymmetric				–

**Table 3** *p* values of SRT for right stimuli

Right	Free ear	Omnidirectional	Directional	Asymmetric
Free ear	–	<0.05*	<0.001***	<0.001***
Omnidirectional		–	<0.001***	<0.01**
Directional			–	<0.001***
Asymmetric				–

**Table 4** *p* values of SRT for back stimuli

Back	Free ear	Omnidirectional	Directional	Asymmetric
Free ear	–	<0.001***	<0.001***	NS
Omnidirectional		–	<0.001***	<0.001***
Directional			–	<0.001***
Asymmetric				–

values < 0.05 in Tables 2, 3, 4, and 5. We then depict, for each direction, the SRT distributions of the four listening conditions in Fig. 6. and we report in a single diagram only the means of the SRT distributions in all directions and condition in Fig. 7.

**Table 5** *p* values of SRT for left stimuli

Left	Free ear	Omnidirectional	Directional	Asymmetric
Free ear	–	<0.01**	<0.001***	<0.001***
Omnidirectional		–	<0.001***	<0.001***
Directional			–	NS
Asymmetric				–

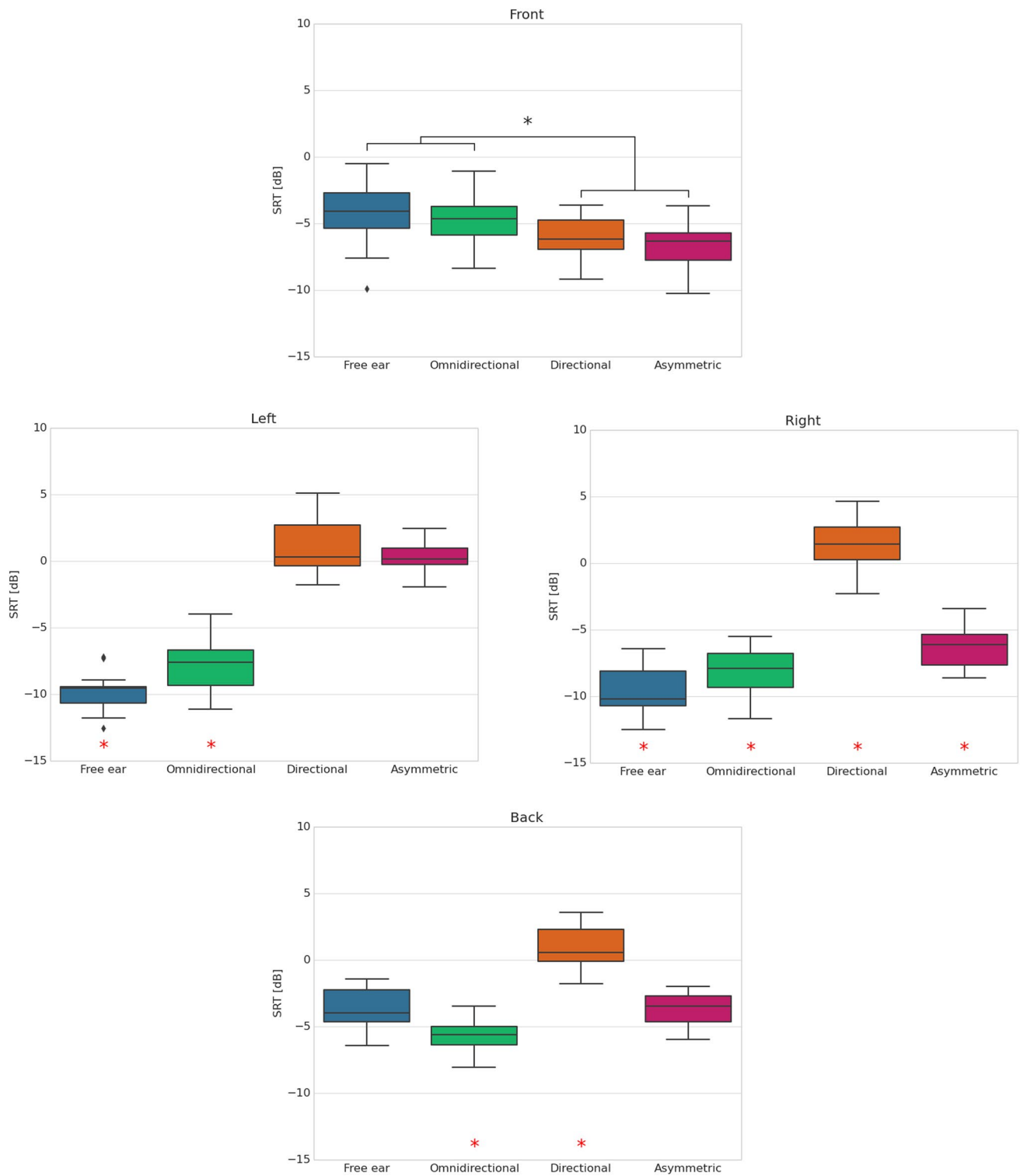
### 3 Results

#### 3.1 Speech reception in *Free ear* follows previous literature. *Omnidirectional* SRT values decrease for frontal and posterior sources, but increase for lateral stimuli

From Fig. 7, it is apparent that the *Free ear* listening condition exhibits the lowest SRT values in correspondence of the lateral stimuli (− 9.84 and − 9.56 dB), followed by the frontal one (− 4.25 dB). The listening pattern of *Free ear* of Fig. 7, in fact, is shaped as a rhomboid. The stimuli coming from the back gave the worst (highest) SRT (− 3.61 dB). The result is aligned with previous literature (Tonning 1971). Therefore, the ability to solve the cocktail party noise is as much accurate as the source is located to the side.

*Omnidirectional* qualitatively exhibits a pattern similar to *Free ear*, with the best SRT for lateral target speech (− 8.10 and − 8.02 dB) and worse values for frontal (− 4.73 dB) and back target speech (− 5.71 dB). Nonetheless, in this condition, SRT values seem to be more balanced, with a lower difference between lateral and the front/back target speech. However, *Omnidirectional* was non-significantly different than *Free ear* only in Front. Instead, it performed better than *Free ear* only in Back (see Fig. 7) and worse than *Free ear* for lateral target speech sources (see Tables 2, 3). The gain of omnidirectional listening is therefore only confined to sources coming from the back.

Finally, we point out the *Free ear* condition had full bandwidth, in that it allowed the participants to perceive

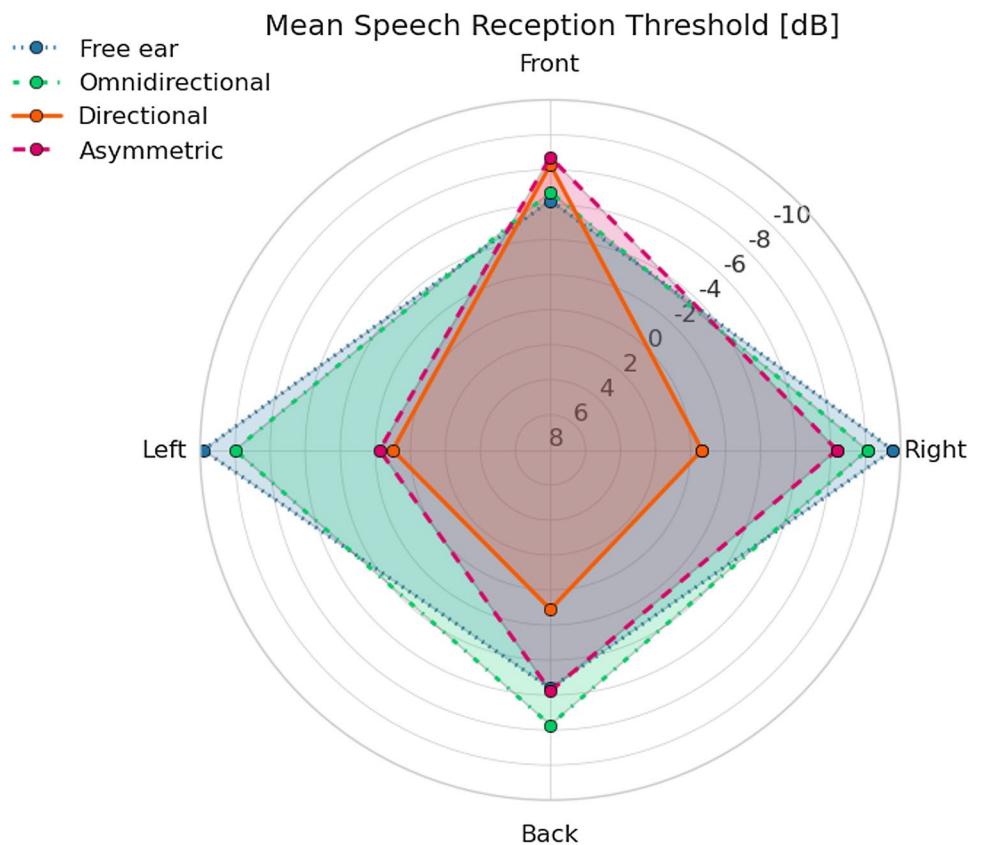


**Fig. 6** Boxplot of SRT values for each listening condition in the different directions. In absolute terms, a SRT distribution that has a median at  $X$  SNR means that it is possible to understand 50% of the sentences up to a SNR of  $X$  dB. Below 0 dB, the energy of the cock-

tail party noise supersedes that of the target speech. Therefore, when comparing SRT distributions, the SRT at lower SNR is more robust to cocktail party noise. Red stars below the boxes indicate that the condition resulted significantly different from all others in the plot



**Fig. 7** Directional listening pattern: mean Speech Reception Thresholds of the participants in four listening conditions. A value farther away from the center means lower Speech Reception Threshold, i.e. better speech understanding in noise



the whole frequency content of the stimuli up to 20 kHz. On one side this represents an ecological baseline, since it reflects the bandwidth the participants are usually exposed to. However, the SRTs of this condition may be different than those obtained by reducing the bandwidth to 8 kHz, as done through the Glassense in the other three conditions, because of the 16 kHz sampling rate of the glasses. It is entirely possible that decreasing the frequency content of the target speech (and of the cocktail party noise) the SRTs in *Free ear* may increase, i.e. be worse.

### 3.2 Directional is better than Free ear and Omnidirectional for frontal sources, as expected

As expected from past literature (Froehlich et al. 2015; Giuliani et al. 2017, 2016; Hornsby and Ricketts 2007; Mens 2011), *Directional* condition showed a significant increase in frontal speech comprehension ( $-6.28$  dB) as compared to *Omnidirectional* and *Free ear*, at the cost of a strong increase of SRT values for lateral (1.34 and 0.99 dB) and posterior (0.90 dB) directions. The improvement of 2 dB of *Directional* compared to *Omnidirectional* condition is lower than those obtained in other studies (Giuliani et al. 2016, 2017). A first possible explanation is the different mean age of the participants (26 in this study, 41.37 in Giuliani et al.

(2016), 72 in Giuliani et al. (2017)), since it has been stated in previous literature that there is a relation between age and speech comprehension in noise for normal hearing and hearing impaired people (Plomp and Mimpen 1979). A second explanation is that, differently from what have been done in other studies, in this setup we decided to use short sentences instead of single words as stimuli. This decision has been taken in order to use a more ecological setup and could have a role in the different contribution of directional advantage in noise. These two possible causes could have contributed, together, to limit the (still significant) amount of improvement of the directional vs omnidirectional condition.

### 3.3 Asymmetric condition similar to Directional for frontal sources, but significantly better than Omnidirectional and Free ear

The *Asymmetric* condition resulted similar to *Directional* (see Fig. 6, 'Front' panel) as for frontal SRT ( $-6.73$  dB, no significant difference), in agreement with (Kim and Bryan 2011; Mens 2011; Picinali and Prosser 2010), but in contrast with other authors (Devocht et al. 2016; Hornsby and Ricketts 2007; Ricketts and Picou 2013). These last studies, unlike our non-significant difference, found the *Asymmetric* condition worse than the *Directional*.

We interpret that *Directional* is similar to *Asymmetric* because of binaural loudness summation, that positively affects both conditions. In particular, the binaural gain in listening the frontal target speech source with both ears may have been higher than the stronger presence of noise in the asymmetric condition. In fact, while the left ear perceives mostly frontal target speech and frontal noise in both conditions, the right ear perceives frontal target speech in both conditions, but a higher noise power in the Asymmetric condition (because the right ear captures noise from all directions). In other words, the omnidirectional noise affecting the right ear may have been masked by a much better reception of the frontal target speech source, which arrived at both ears in both conditions. Binaural summation therefore may have reduced the difference between the two conditions.

### 3.4 *Asymmetric condition better than Directional for speech coming from the omnidirectional filtering side, but worse than Omnidirectional and Free ear*

As expected, *Asymmetric* performed worse than *Omnidirectional* and *Free ear* when target speech came from the left side (the SRT was 0.24 dB), i.e. to the side where the filtering was directional (see Fig. 6, ‘Left’ panel). It did not differ from *Directional*, with a small but non-significant improvement of 0.75 dB. The reason of this result can be different than the previous case, since here binaural summation may have minimally occurred (the target source does not come from the front). In fact, even if in the asymmetric fitting the right ear can perceive the target signal coming from the left side, the target speech power at the right ear could have been reduced by head shadowing (MacKeith and Coles 1971). In other words, hearing the target speech mainly contralaterally in cocktail party noise seems not to help.

A different result was instead obtained when the target speech came from the omnidirectional side (see Fig. 6, ‘Right’ panel). Here the *Asymmetric* (SRT of  $-6.37$  dB) superseded *Directional* (see Tables 1, 3) but was worse than *Omnidirectional* ( $-8.10$  dB), contradicting our initial hypothesis of a similar or better performance of asymmetric fitting with respect to the omnidirectional one.

The explanation of why *Omnidirectional* is better can lie again in binaural summation, mainly occurring in this listening condition, in addition with the fact that lateral sound sources are perceived better in noise (just like the romboid-like listening pattern in *Free ear*, see Fig. 7). Unlike the situation in which the target speech comes from the left (see previous paragraph, SRT = 0.24), here the *Asymmetrical* benefits from a sound target source that is ipsilateral (SRT =  $-6.37$ ): the gain in SRT is remarkable and could mainly be due to absence of head shadowing. Here the right target speech source directly reached the omnidirectional right

ear, while in the previous condition the left target speech source indirectly reached the right omnidirectional ear. We speculate that the contribution of the directional ear did not account for this major change in SRT.

However, our results do not follow those found by Picinali and Prosser. It is possible that the large advantage given by asymmetric configuration with respect to symmetric directional and omnidirectional for lateral sources apparent in their study (not found in our results) was accounted for by our quite different setups. While Picinali and Prosser adopted a simulated cardioid microphone, with binaural listening coming from recordings and no interaural distance, we used a microphone array in an ecological setup. Therefore, our results are affected by natural binaural cues, head shadowing and a different (non-ideal) filtering beam pattern. Picinali and Prosser hypothesized that the particular result they found was due to the activation of the Binaural Masking Level Difference phenomenon, caused by the particular conditions of their simulated setup. The absence of the effect in this work, thus, does not conflict with their hypothesis.

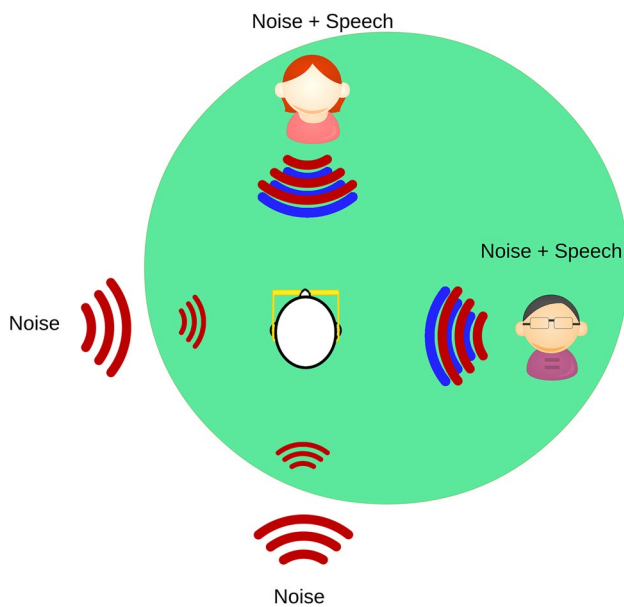
### 3.5 *Asymmetric condition better than Directional and worse than Omnidirectional for speech coming from the back direction*

The comparison between *Asymmetric* and the other conditions showed, finally, that there is a significant advantage of 4.7 dB with respect to *Directional* for posterior sources (see Fig. 6, ‘Back’ panel). Nonetheless, the best condition for this direction is the *Omnidirectional* (1.94 dB of improvement with respect to the *Asymmetric*).

The heavy attenuation of the posterior target speech clearly explains the ineffectiveness of the symmetric directional fitting condition. In fact, the beampattern in Fig. 2 shows, at  $-90^\circ$  (i.e. sound sources coming from the back) attenuations between  $-5$  and  $-15$  dB across the frequency range of interest. In the *Omnidirectional* condition, on the other side, the speech signal is diotically perceived and may activate the binaural loudness summation phenomenon, which increase the resulting SRT. The asymmetric fitting, in which the speech is only perceived monaurally on the right ear, may have inhibited the summation, possibly explaining the lower performance.

## 4 Conclusions

The results obtained in this study indicate that symmetric directional binaural listening, as opposed to omnidirectional, significantly increases speech comprehension of frontal sources in cocktail party noise for healthy hearing people. Moreover, the advantage given by an asymmetric binaural



**Fig. 8** The Asymmetric fitting potentially allows to understand two speech sources disposed at  $90^\circ$  and to perceive competing noise sources from the other two sides as if they had smaller power

configuration does not significantly differ from the symmetric directional condition for frontal speech sources.

However, when the source of interest comes from the side, our SRT values for lateral sources in asymmetric directional listening do not confirm the results obtained in (Picinali and Prosser 2010). The asymmetric improves over the directional listening only for target speech sources ipsilateral to the omnidirectional ear, while no improvement exists for sources of interest contralateral to the omnidirectional ear. Omnidirectional is, anyway, the best listening condition for lateral sources, significantly better than all others, even if less efficient than free ear listening, at least for healthy hearing participants.

In conclusion, asymmetric binaural listening showed to be useful for understanding speech in noise for healthy hearing people. The advantages of this condition are that it is always at least as good as directional listening, with better performance when target sources come from the back. In practice, the asymmetric fitting reduces the disadvantages of directional fittings.

It is possible, even if it has not been tested in this study, that the asymmetric fitting could be effective in facilitating conversations with two non co-located speakers in noisy environments (see Fig. 8). In this kind of situation, in fact, the asymmetric fitting could allow to hear with similar speech comprehension levels the voices of two people disposed at  $90^\circ$ , especially if they are not talking at the same time (a situation which resembles the one tested in this work). This kind of scenario could be attractive not only for people with hearing impairment, but also for people with

normal hearing in particularly noisy environments, like live concerts, crowded demonstrations, where social interaction with multiple talkers is often desired, but difficult (such as requests of help, emergency situations). Admittedly, additional sensors (e.g. cameras, now commonly installed on smart glasses) or sound processing algorithms would be needed to choose which is the principal talker (to the same side of the directional ear) and which is the additional talker (to the same side of the omnidirectional ear): possible solutions include estimating the direction of arrival of the two talkers.

An alternative possible use case could be that in which there is a noise source located in a specific direction that is intended to be excluded, such as working places with running machinery, noisy home appliances or undesired/disturbing human talkers (they correspond to sources from the left in Fig. 8). Here, the asymmetric fitting may activate the directional ear ipsilaterally to the noise source, while maintaining an omnidirectional configuration with the ear contralateral to the noise source. The joint use of direction of arrival estimation algorithms and spectral sound classification could allow to find and isolate specific unwanted acoustic sources without impeding communication in collaborative or social environments.

**Acknowledgements** This research was funded by the Italian Institute of Technology. It received no specific grant from any funding agency in the commercial sectors. The authors would like to thank Giorgio Zini and Francesco Diotalevi from the Electronic Design Lab of the Italian Institute of Technology.

## Compliance with ethical standards

**Conflict of interest** The authors declare that there is no conflict of interest.

**Ethical approval** Written Informed consent was signed by the participants of the study, according to the declaration of Helsinki and the protocol (code 514REG2016) was approved by the local Ethical Committee (ASL3, Regione Liguria, Italy).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Adiloğlu, K., Kayser, H., Baumgärtel, R. M., Rennebeck, S., Dietz, M., & Hohmann, V. (2015). A binaural steering beamformer system for enhancing a moving speech source. *Trends in Hearing*. <https://doi.org/10.1177/2331216515618903>.

- Bocca, E., & Pellegrini, A. (1950). Studio statistico sulla composizione fonetica della lingua italiana e sua applicazione pratica all'audiometria con la parola. *Archivio Italiano Di Otolgia, Rinologia e Laringologia*, 56(5), 116–141.
- Brandstein, M. S. (2000). Microphone arrays: Signal processing techniques and applications. *IEEE Signal Processing Magazine*. <https://doi.org/10.1109/79.826412>.
- Brayda, L., Traverso, F., Giuliani, L., Diotalevi, F., Repetto, S., Sansalone, S., ... Sandini, G. (2015). Spatially selective binaural hearing aids. In: *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2015 ACM international symposium on wearable computers—UbiComp'15*, pp. 957–962. <https://doi.org/10.1145/2800835.2806207>.
- Brimijoin, W. O., McShefferty, D., & Akeroyd, M. A. (2010). Auditory and visual orienting responses in listeners with and without hearing-impairment. *The Journal of the Acoustical Society of America*, 127(6), 3678–3688. <https://doi.org/10.1121/1.3409488>.
- Canzi, P., Manfrin, M., Locatelli, G., Nopp, P., Perotti, M., & Benazzo, M. (2016). Development of a novel Italian speech-in-noise test using a roving-level adaptive method: Adult population-based normative data. *Acta Otorhinolaryngologica Italica*. <https://doi.org/10.14639/0392-100X-1133>.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, 25(5), 975–979. <https://doi.org/10.1121/1.1907229>.
- Chung, K. (2004). Challenges and recent developments in hearing aids: Part I. Speech understanding in noise, microphone technologies and noise reduction algorithms. *Trends in Amplification*. <https://doi.org/10.1177/108471380400800302>.
- Cutugno, F., Prosser, S., & Turrini, M. (2005). *Audiometria Vocale*, Vol III. Padova: GN Resound.
- Devocht, E. M. J., Janssen, A. M. L., Chalupper, J., Stokroos, R. J., & George, E. L. J. (2016). Monaural beamforming in bimodal cochlear implant users: Effect of (A)symmetric directivity and noise type. *PLoS ONE*, 11(8), 1–19. <https://doi.org/10.1371/journal.pone.0160829>.
- Doclo, S., Gannot, S., Moonen, M., & Spriet, A. (2010). Acoustic beamforming for hearing aid applications. In *Handbook on array processing and sensor networks*. <https://doi.org/10.1002/9780470487068.ch9>.
- Doclo, S., & Moonen, M. (2003). Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics. *IEEE Transactions on Signal Processing*, 51(10), 2511–2526. <https://doi.org/10.1109/TSP.2003.816885>.
- Doclo, S., Moonen, M., Van Den Bogaert, T., & Wouters, J. (2009). Reduced-bandwidth and distributed MWF-based noise reduction algorithms for binaural hearing aids. *IEEE Transactions on Audio, Speech and Language Processing*, 17(1), 38–51. <https://doi.org/10.1109/TASL.2008.2004291>.
- European Hearing Instrument Manufacturers Association. (2015). *Trends derived from the EuroTrak databases 2009–2015*.
- European Hearing Instrument Manufacturers Association. (2016). *Hearing Aids improve Hearing—and a LOT more*. Retrieved from [https://www.ehima.com/wp-content/uploads/2016/05/eurotrak-trends-2009-2015-and-mt9\\_.pdf](https://www.ehima.com/wp-content/uploads/2016/05/eurotrak-trends-2009-2015-and-mt9_.pdf).
- Froehlich, M., Freels, K., & Powers, T. A. (2015). Speech recognition benefit obtained from binaural beamforming hearing aids: Comparison to omnidirectional and individuals with normal hearing. *Audiology Online*, (14338), 1–8.
- Gerzon, M. (1974). Surround-sound psychoacoustics. *Wireless World* 80.
- Gerzon, M. A. (1973). Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society*.
- Giuliani, L., Brayda, L., Sansalone, S., Repetto, S., & Ricchetti, M. (2017). Evaluation of a complementary hearing aid for spatial sound segregation. In *IEEE international conference on acoustics, speech and signal processing*. <https://doi.org/10.1109/ICASSP.2017.7952150>.
- Giuliani, L., Sansalone, S., Repetto, S., Traverso, F., & Brayda, L. (2016). Compensating cocktail party noise with binaural spatial segregation on a novel device targeting partial hearing loss. In *International conference on computers helping people with special needs* (pp. 384–391). Springer. [https://doi.org/10.1007/978-3-319-41267-2\\_54](https://doi.org/10.1007/978-3-319-41267-2_54).
- Hornsby, B. W. Y., & Ricketts, T. a. (2007). Effects of noise source configuration on directional benefit using symmetric and asymmetric directional hearing aid fittings. *Ear and Hearing*, 28(2), 177–186. <https://doi.org/10.1097/AUD.0b013e3180312639>.
- Kamkar-Parsi, H., Fischer, E., & Aubreville, M. (2014). New binaural strategies for enhanced hearing. *The Hearing Review*, 21, 42.
- Killion, M. C., & Mueller, H. G. (2010). Twenty years later: A NEW count-the-dots method. *Hearing Journal*. <https://doi.org/10.1097/01.HJ.0000366911.63043.16>.
- Kim, J. S., & Bryan, M. F. (2011). The effects of asymmetric directional microphone fittings on acceptance of background noise. *International Journal of Audiology*, 50(5), 290–296. <https://doi.org/10.3109/14992027.2010.551786>.
- Klasen, T. J., Van den Bogaert, T., Moonen, M., & Wouters, J. (2007). Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues. *IEEE Transactions on Signal Processing*. <https://doi.org/10.1109/TSP.2006.888897>.
- Luts, H., Eneman, K., Wouters, J., Schulte, M., Vormann, M., Buechler, M., ... Spriet, A. (2010). Multicenter evaluation of signal enhancement algorithms for hearing aids. *Journal of the Acoustical Society of America*. <https://doi.org/10.1121/1.3299168>.
- MacKeith, N. W., & Coles, R. R. A. (1971). Binaural advantages in hearing of speech. *The Journal of Laryngology & Otology*. <https://doi.org/10.1017/S0022215100073369>.
- McCormack, A., & Fortnum, H. (2013). Why do people fitted with hearing aids not wear them? *International Journal of Audiology*. <https://doi.org/10.3109/14992027.2013.769066>.
- Mens, L. H. M. (2011). Speech understanding in noise with an eyeglass hearing aid: Asymmetric fitting and the head shadow benefit of anterior microphones. *International Journal of Audiology*, 50(1), 27–33. <https://doi.org/10.3109/14992027.2010.521199>.
- Picinali, L., & Prosser, S. (2010). Monolateral and bilateral fitting with different hearing aids directional configurations. In *Proceedings of the 128th convention of the audio engineering society*.
- Plomp, R., & Mimpen, A. M. (1979). Speech-reception threshold for sentences as a function of age and noise level. *The Journal of the Acoustical Society of America*, 66(5), 1333–1342. <https://doi.org/10.1121/1.383554>.
- Ricketts, T., & Picou, E. M. (2013). Speech recognition for bilaterally asymmetric and symmetric hearing aid microphone modes in simulated classroom environments. *Ear and Hearing*, 34(5), 601–609. <https://doi.org/10.1097/AUD.0b013e3182886d1e>.
- Tonning, F. (1971). Directional audiometry II. The influence of azimuth on the perception of speech. *Acta Oto-Laryngologica*, 72(April), 404–412. <https://doi.org/10.3109/00016487209128420>.
- WHO. (2016). Deafness and hearing loss. *Fact Sheet*. World Health Organization.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.