



## Research Paper

# On the importance of interaural noise coherence and the medial olivocochlear reflex for binaural unmasking in free-field listening

Miriam I. Marrufo-Pérez<sup>a,b</sup>, Leire Araquistain-Serrat<sup>a</sup>, Almudena Eustaquio-Martín<sup>a,b</sup>, Enrique A. Lopez-Poveda<sup>a,b,c,\*</sup>

<sup>a</sup> Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Calle Pintor Fernando Gallego 1, Salamanca 37007, Spain

<sup>b</sup> Instituto de Investigación Biomédica de Salamanca, Universidad de Salamanca, Salamanca 37007, Spain

<sup>c</sup> Departamento de Cirugía, Facultad de Medicina, Universidad de Salamanca, Salamanca 37007, Spain



## ARTICLE INFO

## Article history:

Received 18 November 2020

Revised 25 March 2021

Accepted 31 March 2021

Available online 9 April 2021

## Keywords:

Binaural hearing

Binaural advantage

Speech intelligibility

Hearing device

Noise cancellation

## ABSTRACT

For speech in competition with a noise source in the free field, normal-hearing (NH) listeners recognize speech better when listening binaurally than when listening monaurally with the ear that has the better acoustic signal-to-noise ratio (SNR). This benefit from listening binaurally is known as binaural unmasking and indicates that the brain combines information from the two ears to improve intelligibility. Here, we address three questions pertaining to binaural unmasking for NH listeners. First, we investigate if binaural unmasking results from combining the speech and/or the noise from the two ears. In a simulated acoustic free field with speech and noise sources at 0° and 270° azimuth, respectively, we found comparable unmasking regardless of whether the speech was present or absent in the ear with the worse SNR. This indicates that binaural unmasking probably involves combining only the noise at the two ears. Second, we investigate if having binaurally coherent location cues for the noise signal is sufficient for binaural unmasking to occur. We found no unmasking when location cues were coherent but noise signals were generated incoherent or were processed unilaterally through a hearing aid with linear, minimal amplification. This indicates that binaural unmasking requires interaurally coherent noise signals, source location cues, and processing. Third, we investigate if the hypothesized antimasking benefits of the medial olivocochlear reflex (MOCR) contribute to binaural unmasking. We found comparable unmasking regardless of whether speech tokens (words) were sufficiently delayed from the noise onset to fully activate the MOCR or not. Moreover, unmasking was absent when the noise was binaurally incoherent whereas the physiological antimasking effects of the MOCR are similar for coherent and incoherent noises. This indicates that the MOCR is unlikely involved in binaural unmasking.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## 1. Introduction

Understanding someone's speech while someone else is talking or a machine is buzzing at the same time is a hard task. Fortunately, we have two ears and the head produces an acoustic shadow that causes the signal-to-noise ratio (SNR) to be better in one of the two ears (Kuhn, 1987), be it the ear closer to the target sound source or the ear further away from the interfering sound source. Our brain almost certainly performs the task by focusing on the ear with the better SNR. The advantage of having two ears, however, goes beyond granting an acoustically better ear.

The SNR at which listeners understand 50% of the speech uttered by the target speaker (referred to as the speech reception threshold –SRT– in noise) is 2 to 6 dB better (lower) when listening with two ears than when listening with the acoustically better ear alone (Bronkhorst and Plomp, 1988; Dieudonné and Francart, 2018; Garadat et al., 2009; Litovsky et al., 2017). This benefit from listening binaurally versus listening with the acoustically better ear is known as binaural unmasking (or binaural 'squelch') and, although it has been known for decades (Koenig, 1950), the phenomenon is still poorly understood.

One specific aspect that remains unclear is how the acoustically worse ear contributes to binaural unmasking. There are two possibilities: (1) that it captures some speech, however little it may be, which adds somehow to the speech captured by the better ear;

\* Corresponding author.

E-mail address: [eaopezpoveda@usal.es](mailto:eaopezpoveda@usal.es) (E.A. Lopez-Poveda).

and/or (2) that the noise captured by the acoustically worse ear serves somehow to improve the neural SNR. For example, the brain could use the noise captured by the worse ear as a kind of template to cancel out the noise captured by the acoustically better ear (Durlach, 1963) (alternative mechanisms are discussed later). Simply put, the ear with the worse acoustic SNR could serve to enhance the speech and/or to attenuate or cancel the noise. The first mechanism would require having the speech in the acoustically worse ear while the second mechanism would require having only the noise. To our knowledge, previous studies have not investigated whether binaural unmasking in free-field listening requires only the speech, only the noise, or both the speech and the noise to be present in the ear with worse acoustic SNR and studies that presented stimuli over headphones are insufficient to address this question. For instance, intelligibility is worse for unilateral speech in unilateral noise than for unilateral speech in bilateral diotic noise (Licklider, 1948), suggesting that the brain requires the noise to be present at two ears to produce a binaural advantage. For bilateral speech and noise, intelligibility is worse when the two stimuli are in phase than when the speech is out of phase and the noise is in phase (Johansson and Arlinger, 2002; Licklider, 1948), suggesting that the brain combines the speech information at the two ears to produce a binaural advantage. Speech intelligibility tests with headphones, however, did not include natural head-related interaural cues, and thus listeners did not have an acoustically better and worse ear. Therefore, it is uncertain if the findings from headphone experiments are representative of the results in free-field listening conditions. The present Experiment 1 is aimed at elucidating if binaural unmasking in simulated free-field listening conditions requires having the speech and/or the noise in the acoustically worse ear.

If binaural unmasking in free-field listening occurred because the brain cancels out the background noise, altering the interaural coherence of the noise should reduce binaural unmasking because the auditory brain would be less efficient at cancelling incoherent than coherent noise (Durlach, 1963). In the most general case, for a noise source in the free field, the noise waveform at the left,  $L(t)$ , and the right,  $R(t)$ , eardrum is each determined by three factors:

$$L(t) = N_L(t) * HRIR_L(t) * D_L(t) \quad (1a)$$

$$R(t) = N_R(t) * HRIR_R(t) * D_R(t) \quad (1b)$$

where  $HRIR_L(t)$  and  $HRIR_R(t)$  are the head-related impulse responses (HRIRs) for the left and the right ear, respectively;  $D_L(t)$  and  $D_R(t)$  denote the effect of processing the noise through a hearing device in the left and right ear, respectively; and  $N_L(t)$  and  $N_R(t)$  denote the waveforms generated by the noise source for the left and the right ear, respectively. Of course, in natural listening situations the noise source generates a common waveform for the two ears [i.e.,  $N_L(t) = N_R(t)$ ], and  $HRIR_L(t)$  and  $HRIR_R(t)$  encompass the head-related interaural time and level differences. Existing experimental evidence shows that binaural unmasking is reduced when head-related interaural time differences (ITDs) but not when interaural level differences (ILDs) in the noise are eliminated (Bronkhorst and Plomp, 1988; Dieudonné and Francart, 2018). In other words, binaurally coherent ITDs cues (pertaining to noise source location) are necessary for binaural unmasking to occur. It remains uncertain, however, if binaurally coherent location cues are sufficient or if the interaural coherence is also required for the noise signals themselves [i.e., coherence between  $N_L(t)$  and  $N_R(t)$ ] and/or for their processing through a hearing device [i.e., coherence between  $D_L(t)$  and  $D_R(t)$ ]. The present Experiments 2 to 4 address these questions.

The auditory system uses information from both ears to produce a binaural advantage, but it does not always integrate bin-

aural information quickly. Binaural unmasking builds up over time with a time constant of a few tens of milliseconds when the noise is identical in both ears (diotic), but of ~200 ms when the interaural parameters of the masker (ITDs or interaural phase differences) are changing rapidly (Grantham and Whightman, 1979; Culling and Summerfield, 1998; Hauth and Brand, 2018). In principle, peripheral mechanisms mediated by the medial olivocochlear reflex could contribute to a slow build-up of binaural unmasking over time. In the mammalian auditory system, efferent fibers project from the medial region of the superior olivary complex to outer hair cells (Warr and Guinan, 1979) and their activation inhibits basilar membrane responses to low and moderate sound levels (Cooper and Guinan, 2006; Murugasu and Russell, 1996). This inhibition increases the dynamic range of auditory nerve fibers for tones in noise (Winslow and Sachs, 1988) and presumably facilitates speech-in-noise recognition (Guinan, 2006; Lopez-Poveda, 2018). Medial olivocochlear (MOC) fibers can be activated reflexively by ipsilateral and contralateral sounds, and the time course of activation of the MOC reflex (MOCR) is around 300 ms (Backus and Guinan, 2006). Therefore, MOCR activation could enhance or contribute to the 'sluggish' component of binaural unmasking in three ways: (1) the stimuli presented in the acoustically worse ear could activate the contralateral MOCR and improve the auditory nerve representation of speech in noise, thus the neural SNR, in the better ear; (2) the stimuli presented in the acoustically better ear could activate the contralateral MOCR and increase the neural SNR in the worse ear; and (3) the stimuli presented in the acoustically worse ear could activate the ipsilateral MOCR and increase the neural SNR in the worse ear.<sup>1</sup> In other words, MOCR activation could result in a better representation of the masked speech in both auditory nerves, which could increase binaural unmasking compared to when the MOCR is not active. In addition, because activation of the contralateral MOCR can enhance auditory nerve responses to tones in noise both when the contralateral noise is correlated or uncorrelated with the masking noise (e.g., Fig. 4 in Kawase and Liberman, 1993), the MOCR should contribute to binaural unmasking both for binaurally coherent and incoherent noises. The present Experiment 5 is aimed at elucidating the potential contribution of the MOCR to binaural unmasking.

## 2. Experiment 1: mechanisms of binaural unmasking

### 2.1. Rationale

The aim of this experiment was to elucidate if binaural unmasking in free-field listening requires only the speech, only the noise, or both the speech and the noise to be present at the two ears.

### 2.2. Methods

**Participants.** Sixteen normal-hearing listeners (3 men; mean  $\pm$  SD age = 24.6  $\pm$  3.1 years) participated in the experiment. Fifteen of them had audiometric thresholds  $\leq$  20 dB hearing level (HL) in both ears at octave frequencies between 250 Hz and 8 kHz (ANSI, 1996). One participant had audiometric thresholds equal to 25 dB HL at 8 kHz in the right ear, and  $\leq$  20 dB HL at other frequencies. All of them were native speakers of Spanish. All participants were volunteers and signed an informed consent before they were admitted to the study. The study was approved by the Ethics Committee of the University of Salamanca (Spain).

<sup>1</sup> Note that the ipsilateral MOCR in the better acoustic ear is not expected to contribute to binaural unmasking. This is because binaural unmasking is the difference between listening with the better acoustic ear alone and listening with the two ears, and the ipsilateral MOCR would be already active when listening with the better ear alone.

**Experimental conditions.** SRTs were measured in simulated free-field listening conditions for a target (female) speaker at 0° azimuth and a noise source at 270° azimuth. Free-field listening was simulated by convolving monophonic stimuli with HRIRs for an acoustic manikin (Gardner and Martin, 1995). SRTs were measured in three listening conditions: listening binaurally ( $S_{Bi}N_{Bi}$ ), listening monaurally with the acoustically better ear alone ( $S_mN_m$ ), and listening with the two ears but removing the speech from the acoustically worse ear (the left ear in this case) before the stimulus was presented to the listener ( $S_mN_{Bi}$ ). If the noise but not the speech captured by the acoustically worse ear is sufficient for binaural unmasking to occur, then SRTs should be equal for  $S_{Bi}N_{Bi}$  and  $S_mN_{Bi}$ .

**Stimuli.** Thirty-five disyllabic words (Cárdenas and Marroero, 1994) were used to measure each SRT. The first 10 words were always the same and were used to familiarize participants with the test condition. The last 25 words corresponded to one of the nine phonetically balanced lists from the test. All words were presented in random order. A steady-state noise with a speech-shaped long-term spectrum was used as a masker. The noise started 50 ms before the word onset and finished 50 ms after the word offset.

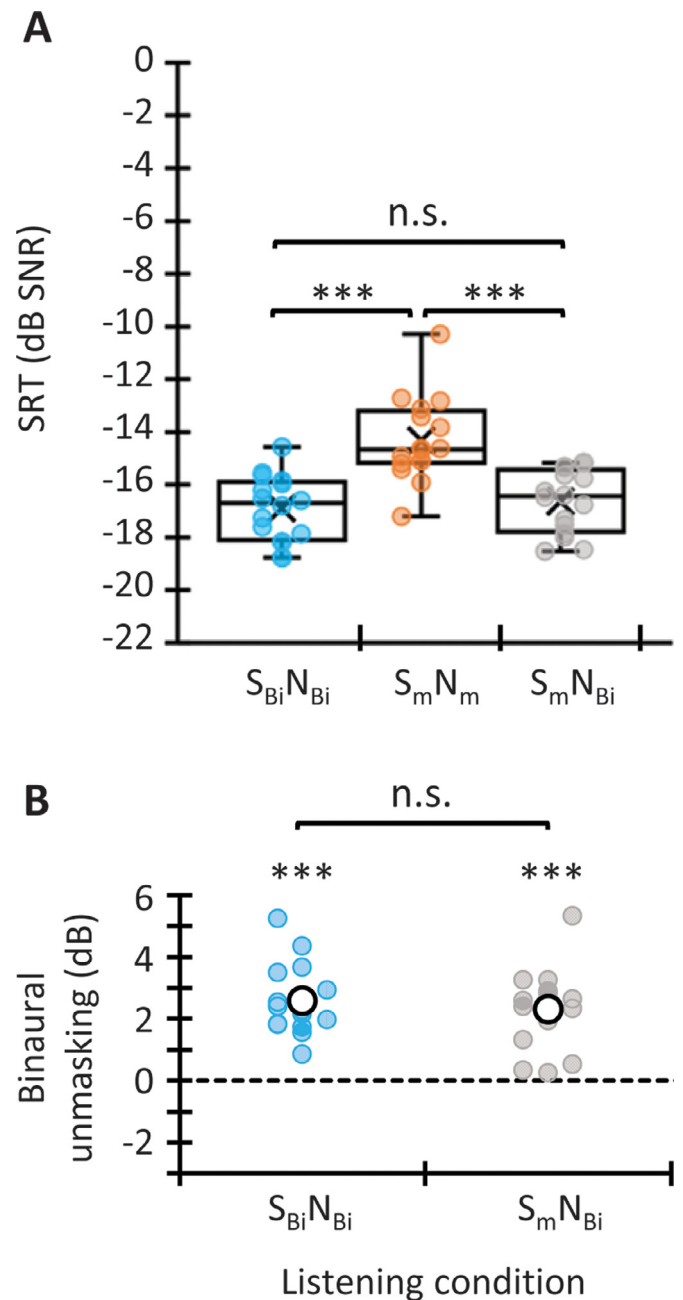
**Procedure.** The speech level was fixed at 65 dB SPL and the noise level increased or decreased adaptively after a correct or an incorrect response, respectively, to obtain the SNR at which listeners recognized 50% of the words uttered by the speaker (Levitt, 1971). The level change was 4 dB between words 1 and 14, and 2 dB between words 14 and 35. The SRT was calculated as the mean SNR for the final 17 words. A sound cue (1 kHz pure tone with 500 ms duration) was presented 500 ms before the stimulus onset to warn the listener about the stimulus presentation and to focus his/her attention on the speech recognition task. Feedback was not given to the participants on the correctness of their responses. Three SRTs were always obtained for each test condition, and their mean was regarded as the SRT. A fourth SRT was measured and included in the mean when the experimenter deemed that the first three SRT measurements were too variable (typically when the across-measures SD  $\geq 1.3$  dB).

**Apparatus.** Stimuli were generated digitally using custom-made software and played via an RME Fireface UCX soundcard at a sampling rate of 44.1 kHz and with 24-bit resolution. Listeners were seated in a double wall sound attenuating booth during all measurements. Stimuli were presented via Etymotic ER-2 insert earphones, which are designed to have a flat frequency response at the eardrum (0.1–16 kHz) and thus to preserve head-related cues. SPLs were calibrated by placing the earphones in a Zwislocki DB-100 coupler connected to the sound level meter (Brüel & Kjaer 2238). Calibration was performed at 1 kHz and the obtained sensitivity was used at all other frequencies.

**Statistical analyses.** Kolmogorov-Smirnov tests revealed that SRTs conformed to a normal distribution. Repeated-measures analysis of the variance (RMANOVA) tests were used to compare mean SRTs and the amount of binaural unmasking across conditions. Post hoc pairwise comparisons were conducted using Bonferroni corrections for multiple comparisons. Two-tailed significance levels were applied for all tests. An effect was regarded as statistically significant when the null hypotheses could be rejected with 95% confidence ( $p \leq 0.05$ ).

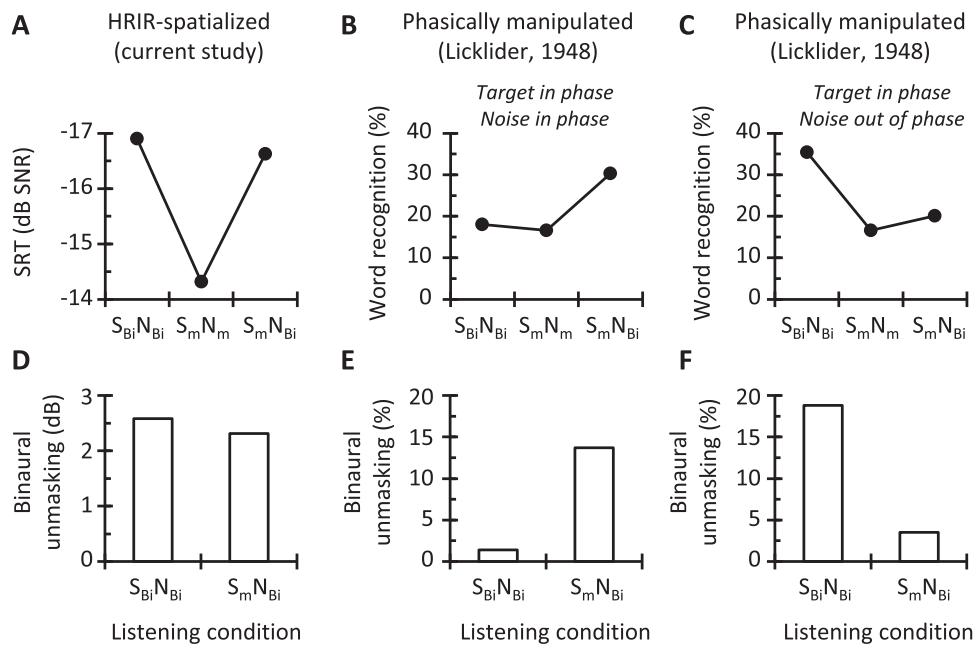
### 2.3. Results

**Fig. 1A** shows SRTs when participants listened binaurally ( $S_{Bi}N_{Bi}$ ), monaurally with the acoustically better ear alone ( $S_mN_m$ ), and with their two ears but the speech was removed from the acoustically worse ear ( $S_mN_{Bi}$ ). The RMANOVA revealed a statistically significant effect of listening condition on SRT [ $N = 16$ ,  $F(2,30)=55.3$ ,  $p<0.001$ ]. Post hoc pairwise comparisons revealed



**Fig. 1.** Results of Experiment 1. A. SRTs for words in noise for 16 NH listeners in three listening conditions: listening binaurally ( $S_{Bi}N_{Bi}$ ), listening monaurally with the acoustically better ear alone ( $S_mN_m$ ), and listening with the two ears but removing the speech from the acoustically worse ear ( $S_mN_{Bi}$ ). The bottom, middle and top lines in each box plot denote the 25, 50 (median), and 75 percentiles, respectively. Crosses illustrate the mean values. Circles show individual scores (mean of three SRT estimates). B. Binaural unmasking calculated as the difference between SRTs in the monaural condition ( $S_mN_m$ ) and in the other conditions. Positive values indicate worse SRTs in the monaural than in the other condition. Filled circles depict individual results and open circles depict mean values. Statistically significant pairwise comparisons are indicated by asterisks: \*\*\* $p \leq 0.001$ . n.s.: not significant.

that SRTs were significantly better (lower) when participants listened binaurally than when they listened with their acoustically better ear alone, demonstrating binaural unmasking ( $S_{Bi}N_{Bi}$  vs  $S_mN_m$  in **Fig. 1A**;  $p<0.001$ ). Interestingly, when the speech but not the noise was removed from the acoustically worse ear ( $S_mN_{Bi}$ ), SRTs were not significantly different from those in the binaural listening condition ( $S_mN_{Bi}$  vs  $S_{Bi}N_{Bi}$  in **Fig. 1A**;  $p = 0.626$ ) but were still significantly better than SRTs in the better-ear listening condi-



**Fig. 2.** Comparison of the present results with those from Licklider (1948). A. Mean SRTs for Experiment 1 of the present study. Note the inverted scale on the y-axis to facilitate the comparison with the results from Licklider (1948). B and C. Percentage of recognized words in white noise masker for a fixed SNR of  $-10$  dB. Results are for NH listeners (Table II of Licklider, 1948). Stimuli were not convolved with HRIRs.  $S_{Bi}N_{Bi}$ : bilateral speech and noise.  $S_mN_m$ : unilateral speech and noise.  $S_mN_{Bi}$ : unilateral speech and bilateral noise. In panel B, the speech and the noise had the same phase across the ears (i.e., they were in phase) when they were bilateral. In panel C, the speech was in phase when it was bilateral, and the noise was out of phase. Note that we chose the conditions where the speech was always in phase because we deemed it the closest condition to our study (our speech source was always in front and in phase across the ears). When the speech and/or noise were unilateral, results were for the right ear (as in our study). D, E, and F. Unmasking calculated as the difference between SRTs in the monaural condition ( $S_mN_m$ ) and in the other conditions of panels A, B, and C, respectively.

tion ( $S_mN_{Bi}$  vs  $S_mN_m$  in Fig. 1A;  $p < 0.001$ ). Mean binaural unmasking, calculated as the difference relative to the monaural condition, was 2.6 dB for the  $S_{Bi}N_{Bi}$  condition and 2.3 dB for the  $S_mN_{Bi}$  condition (Fig. 1B) and was not statistically different for the two conditions [ $t(15) = 1.3$ ;  $p = 0.224$ ].

#### 2.4. Discussion

To our knowledge, this is the first study to show that when the speech and noise sources are spatially separated in the horizontal plane, binaural unmasking is similar when participants are presented with binaural noise and speech as when they are presented with binaural noise and monaural speech in the acoustically better ear. This demonstrates that the speech captured by the acoustically worse ear is dispensable for binaural unmasking to occur.

Previous studies have measured binaural unmasking for conditions that resemble those tested here but without simulating free-field listening (e.g., Licklider, 1948). A comparison of the results across studies suggests that it would not be possible to infer the present findings from existing data (Fig. 2). Fig. 2A shows mean SRTs for words in noise when stimuli were presented over headphones but convolved with HRIRs to simulate free-field listening (replotted from Experiment 1), and Fig. 2B-C show word-in-noise recognition scores for stimuli presented over headphones without being convolved with HRIRs (i.e., free-field listening was not simulated; data from Licklider, 1948). Conditions were the closest as possible: bilateral speech and noise ( $S_{Bi}N_{Bi}$ ), unilateral speech and noise ( $S_mN_m$ ), or unilateral speech in bilateral noise ( $S_mN_{Bi}$ ). Our target speech was always in front (at  $0^\circ$  azimuth) and thus in phase across the ears. For this reason, we chose the results from Licklider (1948) where speech was in phase (panels B and C), and the noise was also in phase (panel B) or out of phase (panel C). Fig. 2D-F depict the amount of binaural unmasking calculated as the difference between the unilateral noise ( $S_mN_m$ ) and bilateral

noise ( $S_{Bi}N_{Bi}$  or  $S_mN_{Bi}$ ) conditions. In free-field listening (present results), binaural unmasking is similar in magnitude when listening to binaural speech in binaural noise ( $S_{Bi}N_{Bi}$ ) as when listening to monaural speech in binaural noise ( $S_mN_{Bi}$ ), i.e., with the speech omitted in the acoustically worse ear (Fig. 2D). Binaural unmasking for stimuli not convolved with HRIRs is similar for unilateral speech in bilateral noise ( $S_mN_{Bi}$ ) as for bilateral speech in bilateral noise ( $S_{Bi}N_{Bi}$ ) but only when the noise is in phase in the former condition and out of phase in the later condition (Fig. 2E and 2F). In other words, in simulated free-field listening, binaural unmasking is similar for the  $S_mN_{Bi}$  and  $S_{Bi}N_{Bi}$  conditions when the interaural phase of the noise is identical in both conditions [the noise source location did not change from one condition to another] (Fig. 2D). However, when stimuli are not convolved with HRIRs, binaural unmasking is different for the  $S_mN_{Bi}$  and  $S_{Bi}N_{Bi}$  conditions if the noise phase is identical in both conditions (Fig. 2E or 2F). Altogether, the present results suggest that the benefit on speech intelligibility of adding a second ear needs not be analogous when it is measured without or with the simulation of free-field listening.

The reasons for the discrepancy across studies are uncertain. There are at least three possibilities. First, it is possible that because in the present experiment listeners had enough cues to perceive the spatial location of the noise source, they could focus on the speech presented in the better ear, thus resulting in similar SRTs in the  $S_mN_{Bi}$  and  $S_{Bi}N_{Bi}$  conditions (but see the General Discussion). In the study of Licklider (1948), by contrast, there was no better or worse acoustic ear, and listeners probably attended to the speech presented in the two ears, which could have been detrimental when the speech and noise were both in phase (Fig. 2B). Second, whereas the locations of the speech and noise sources did not change across conditions in our study, the subjective impression concerning the locations of the speech and noise changed in the study of Licklider (1948) depending on the phase of the speech

and the noise (his Fig. 6). This can make it difficult to integrate the results of the two studies. Third, it is also possible that in our study, the SNR in the acoustically worse ear was so low that listeners could not detect or recognize any speech with that ear, thus resulting in similar SRTs in the  $S_{mN_{Bi}}$  and  $S_{BiN_{Bi}}$  conditions.

### 3. Experiments 2–4: the importance of binaural noise coherence for binaural unmasking

#### 3.1. Rationale

The results from Experiment 1 suggest that binaural unmasking in free-field listening requires only the noise but not the speech to be present in the two ears. In this section, we address a complementary question: how similar (coherent) must the noise in the two ears be for binaural unmasking to occur? We report three experiments, each intended to alter the interaural coherence of the noise in a different way: (1) altering noise source location cues (ITDs and ILDs); (2) altering the interaural coherence of the noise signal itself, i.e., generating different noise signals for the left and the right ears but from the same location; (3) wearing a hearing aid in one of the ears. The aim is to investigate if the interaural cues that determine the spatial location of the noise source are necessary and sufficient for binaural unmasking to occur.

#### 3.2. Methods

*Experimental conditions.* Three experiments were conducted, each with a different number of participants, as described below.

- **Experiment 2.** The aim was to investigate if binaurally coherent noise source location cues (ITDs and ILDs) are necessary for binaural unmasking to occur. SRTs were measured for the three conditions tested in Experiment 1 ( $S_{BiN_{Bi}}$ ,  $S_{mN_m}$ ,  $S_{mN_{Bi}}$ ) plus two other conditions ( $S_{mN_\phi}$ , and  $S_{mN_{Lrov}}$ ). Condition  $S_{mN_\phi}$  was as  $S_{mN_{Bi}}$  except that the noise waveform was inverted (180° phase shift) in the ear with the worse acoustic SNR before it was presented to the listener (i.e., after applying the HRIRs). This phase shift is expected to distort the instantaneous ITDs and ILDs of the noise but not the onset ITDs or the long term ILDs. Condition  $S_{mN_{Lrov}}$  was as  $S_{mN_{Bi}}$  except that the noise level was randomly roved (by up  $\pm 5$  dB) in the ear with the worse acoustic SNR before it was presented to the listener. This would distort the noise ILDs but not ITDs. The presence of normal binaural unmasking in the  $S_{mN_{Lrov}}$  condition with reduced binaural unmasking in the  $S_{mN_\phi}$  condition would suggest that instantaneous ITDs are important for binaural unmasking. According to previous studies, this is expected to occur (Garadat et al., 2009). Participants and free-field listening simulation procedures were as in Experiment 1. The results for  $S_{BiN_{Bi}}$ ,  $S_{mN_m}$ , and  $S_{mN_{Bi}}$  conditions have already been presented in Experiment 1, but they will be presented here again for comparison purposes.

- **Experiment 3.** The aim was to investigate if the noise waveforms generated for the two ears need to be binaurally coherent [i.e., if  $N_L(t)$  must be equal to  $N_R(t)$ , Eq. 1] for binaural unmasking to occur. As in Experiment 1, SRTs in simulated free-field listening were measured for the  $S_{BiN_{Bi}}$ ,  $S_{mN_m}$  and  $S_{mN_{Bi}}$  conditions and a new condition was added ( $S_{mN_{Inc}}$ ) where the noise signals at the two ears were HRIR-convolved to provide coherent location cues but were different (incoherent), as if they had been generated by different sound sources [i.e.,  $N_L(t) \neq N_R(t)$  in Eq. (1)]. If binaural unmasking were small or absent in this new condition ( $S_{mN_{Inc}}$ ), this would suggest that coherent noise source location cues are insufficient for unmasking to occur and that unmasking also requires interaurally coherent noise source signals. Seventeen NH listeners participated in this experiment

(5 men; mean  $\pm$  SD age = 24.0  $\pm$  3.7 years), of whom 11 were common to Experiments 1 and 2. Sixteen participants had audiometric thresholds  $\leq 20$  dB HL in both ears at octave frequencies between 250 Hz and 8 kHz. One participant had audiometric thresholds equal to 25 dB HL at 8 kHz in the right ear. All participants were native speakers of Spanish. For seven participants, SRTs for the  $S_{BiN_{Bi}}$ ,  $S_{mN_m}$  and  $S_{mN_{Bi}}$  conditions were carried forward from Experiment 1.

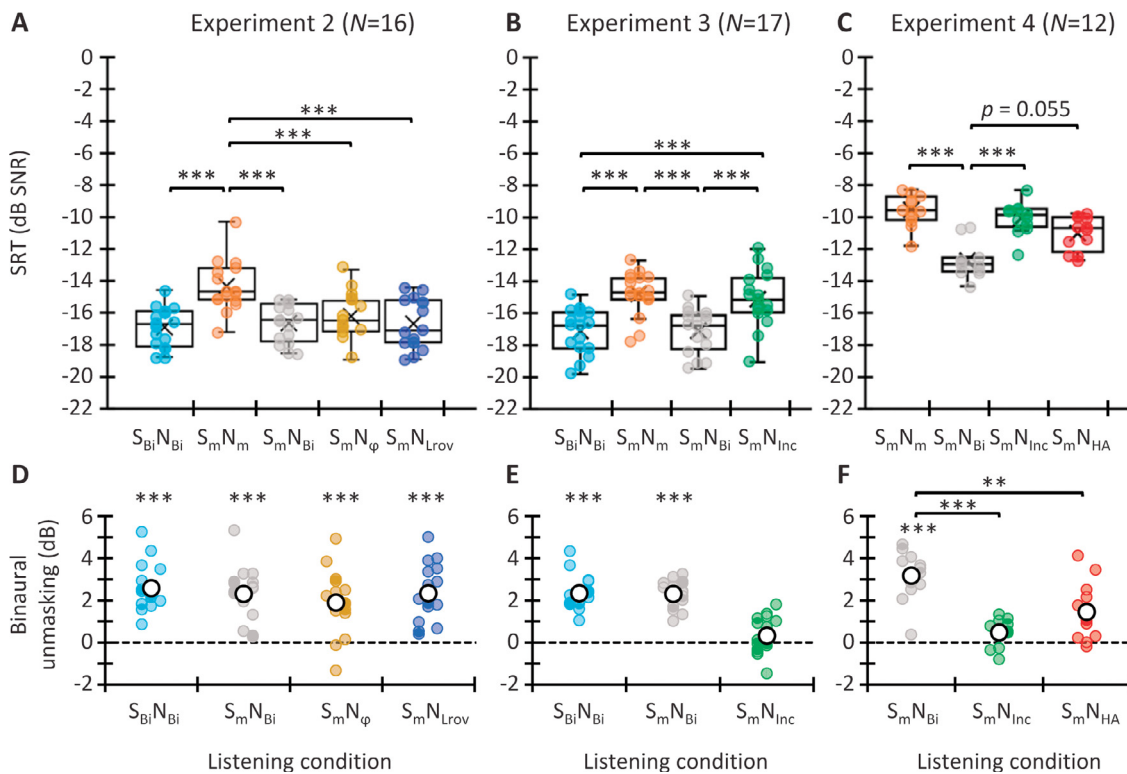
- **Experiment 4.** The aim was to investigate if binaural unmasking decreases or disappears when the binaural coherence of the noise is altered by processing the noise through a hearing aid in one of the ears. SRTs were measured for unilateral speech in unilateral noise ( $S_{mN_m}$ ), unilateral speech in bilateral noise ( $S_{mN_{Bi}}$ ), unilateral speech in bilateral incoherent noise ( $S_{mN_{Inc}}$ ), and unilateral speech in bilateral coherent noise but the listener wore a hearing aid in the ear opposite to the speech ( $S_{mN_{HA}}$ ). We used a Starkey series iQ11 experimental hearing aid fitted with a linear gain  $\leq 5$  dB across the frequency range 125–8000 Hz. Except for multiband processing and feedback cancellation, all other signal processing features in the hearing aid were inactivated for the trial. In this experiment, stimuli were not convolved with HRIRs and so participants did not have an acoustically better or worse ear. Instead, the speech was always presented in the right ear and the noise was presented in the right ear ( $S_{mN_m}$ ) or in both ears ( $S_{mN_{Bi}}$ ,  $S_{mN_{Inc}}$ ,  $S_{mN_{HA}}$ ). Twelve NH listeners participated in the experiment (5 men; mean  $\pm$  SD age = 29.9  $\pm$  7.5 years). Ten of them had audiometric thresholds  $\leq 20$  dB HL at both ears at octave frequencies between 250 Hz and 8 kHz. Two participants had audiometric thresholds equal to 25 dB HL at one frequency in one ear, and  $\leq 20$  dB HL at other frequencies. All but one participant were native speakers of Spanish. The non-native Spanish speaker participant had a high level of proficiency in Spanish.

Stimuli and procedures were as for Experiment 1. Experiments 2 and 3 were carried out with the same equipment as Experiment 1. For Experiment 4, stimuli were presented via circumaural, open, Sennheiser HD580 headphones. In the  $S_{mN_{HA}}$  condition, the hearing aid was located behind the participant's left ear and under the headphone and care was exercised that the headphone earpad did not occlude the hearing aid microphones in the behind-the-ear piece. SPLs were calibrated by placing the headphones on a KEMAR head (Knowles Electronics) equipped with a Zwislocki (Knowles Electronics DB-100) artificial ear connected to a sound level metre (Brüel & Kjaer 2238). Calibration was performed at 1 kHz and the obtained sensitivity was used at all other frequencies.

*Statistical analyses.* Kolmogorov-Smirnov tests revealed that SRTs conformed to a normal distribution for Experiments 2 and 3 but not for Experiment 4. RMANOVA tests were used to compare SRTs and the amount of binaural unmasking across conditions for Experiments 2 and 3, and a Friedman test was used for Experiment 4. Post hoc pairwise comparisons were conducted using Bonferroni corrections for multiple comparisons. Two-tailed significance levels were applied for all tests. Greenhouse-Geisser corrections were applied when the sphericity assumption was not satisfied.

#### 3.3. Results

Experiment 2 was aimed at investigating the effect of altering the noise ILDs and ITDs on binaural unmasking (Fig. 3A and 3D). A RMANOVA revealed a statistically significant effect of listening condition on SRT [ $N = 16$ ,  $F(4,60) = 26.8$ ,  $p < 0.001$ ]. SRTs were significantly higher in the monaural condition than when the noise was presented to the two ears but the waveform was inverted in the acoustically worse ear ( $S_{mN_m}$  vs  $S_{mN_\phi}$  in Fig. 3A,  $p = 0.001$ ). SRTs were also significantly higher in the monaural



**Fig. 3.** Results of Experiments 2–4. A. SRTs for Experiment 2.  $S_{mN_{\phi}}$  refers to monaural speech and binaural noise with the noise phase inverted in the worse ear.  $S_{mN_{Lrov}}$  refers to monaural speech and binaural noise but with the noise level roved by up  $\pm 5$  dB in the worse ear. B. SRTs for Experiment 3.  $S_{mN_{Inc}}$  refers to monaural speech and binaural incoherent noise. C. SRTs for Experiment 4. Stimuli were not convolved with HRIRs.  $S_{mN_{HA}}$  refers to unilateral speech and bilateral coherent (diotic) noise but the listener wore a hearing aid in the ear opposite from the speech. D, E, F. Binaural unmasking for Experiments 2, 3, and 4, respectively. It was calculated as the difference between SRTs in the monaural condition ( $S_{mN_m}$ ) relative to the other test conditions. The layout is as for Fig. 1. Statistically significant pairwise comparisons are indicated by asterisks:  $**p \leq 0.01$ ;  $***p \leq 0.001$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

condition than when the noise was presented to the two ears and the noise level was roved in the worse ear ( $S_{mN_m}$  vs  $S_{mN_{Lrov}}$  in Fig. 3A,  $p < 0.001$ ). In other words, binaural unmasking occurred in all conditions where the noise was presented to the two ears. Binaural unmasking was not significantly different across the  $S_{BiN_{Bi}}$ ,  $S_{mN_{Bi}}$ ,  $S_{mN_{\phi}}$  and  $S_{mN_{Lrov}}$  conditions [ $F(3,45)=2.3$ ,  $p = 0.088$ ], but it tended to be smaller when inverting the noise waveform than when roving the noise level (1.9 vs 2.4 dB, respectively; Fig. 3D).

Experiment 3 was aimed at investigating the effect of interaural noise signal coherence on binaural unmasking (Fig. 3B and 3E). We found a statistically significant effect of listening condition on SRT [ $N = 17$ ,  $F(1.9,30.1)=60.7$ ,  $p < 0.001$ ]. SRTs were better in the  $S_{BiN_{Bi}}$  or  $S_{mN_{Bi}}$  conditions than in the  $S_{mN_m}$  condition, consistent with the results of Experiment 1. The new finding is that binaural unmasking disappeared when the noise signals were HRIR-convolved to provide binaurally coherent location cues but were generated incoherent ( $S_{mN_m}$  vs  $S_{mN_{Inc}}$  in Fig. 3B;  $p = 0.642$ ). Binaural unmasking for incoherent noise signals was 0.3 dB and was not statistically different from zero (Fig. 3E).

Experiment 4 was aimed at investigating the effect of hearing device processing on binaural unmasking when the auditory system is not damaged (Fig. 3C and 3F). We found a statistically significant effect of listening condition on SRT [ $N = 12$ ,  $\chi^2(3)=27.5$ ,  $p < 0.001$ ]. Post-hoc comparisons revealed that SRTs were significantly better when the speech was unilateral and the noise was bilateral and coherent (diotic) than when the speech and noise were unilateral ( $S_{mN_m}$  vs  $S_{mN_{Bi}}$  in Fig. 3C;  $\chi^2=2.6$ ,  $p < 0.001$ ). This binaural advantage did not reach statistical significance when the noise was binaurally incoherent (dichotic) ( $S_{mN_m}$  vs  $S_{mN_{Inc}}$  in Fig. 3C;  $\chi^2=0.6$ ,  $p = 1.000$ ) or was binaurally coherent but the listener wore a hearing aid in the ear opposite from the speech

( $S_{mN_m}$  vs  $S_{mN_{HA}}$  in Fig. 3C;  $\chi^2=1.3$ ,  $p = 0.106$ ). SRTs in the (unilateral) hearing-aid condition were close to SRTs in the condition where the noise was binaurally incoherent ( $S_{mN_{HA}}$  vs  $S_{mN_{Inc}}$  in Fig. 3C;  $\chi^2=-0.6$ ,  $p = 1.000$ ). Binaural unmasking was 3.2, 0.5, and 1.5 dB for the  $S_{mN_{Bi}}$ ,  $S_{mN_{Inc}}$ , and  $S_{mN_{HA}}$  conditions, respectively (Fig. 3F). A Friedman test revealed differences in the amount of binaural unmasking across listening conditions [ $\chi^2(2)=19.5$ ,  $p < 0.001$ ]. Binaural unmasking was greater when the noise was coherent ( $S_{mN_{Bi}}$ ) than when it was incoherent ( $S_{mN_{Inc}}$ ) ( $p < 0.001$ ) or than when listeners wore a hearing aid ( $S_{mN_{HA}}$ ) ( $p < 0.007$ ) (Fig. 3F).

### 3.4. Discussion

In Experiment 2 we investigated the importance of binaurally coherent location cues on binaural unmasking. For simulated free-field listening conditions, binaural unmasking occurred when the noise waveform in the acoustically worse ear was inverted ( $180^\circ$  phase shift) before it was presented to the listener, although the magnitude of unmasking in this condition tended to be smaller than when the noise level was roved in the worse ear. This result is consistent with previous studies that reported a reduction in binaural unmasking after removing ITDs in the masker envelope and/or temporal fine structure (TFS) without changing the coherence of the masker source (Bronkhorst and Plomp, 1988; Dieudonné and Francart, 2018; Garadat et al., 2009). On the other hand, we observed that roving the level of the noise in the worse ear up to  $\pm 5$  dB did not affect binaural unmasking, thus suggesting that binaurally coherent ITDs in the masker are not necessary for binaural unmasking to occur. This result is also consistent with previous studies. Arsenault and Punch (1999) and Bronkhorst and

Plomp (1988) attenuated 20 dB the noise and speech in the acoustically worse ear and did not observe an effect on the magnitude of unmasking. Dieudonné and Francart (2018) demonstrated greater binaural unmasking for a speech-shaped noise when only ITDs were present in the noise than when ITDs and ILDs were present (4.2 vs 2.5 dB, respectively), suggesting that ILDs are dispensable to observe binaural squelch.

We found that binaural unmasking disappeared when the noise maintained location-related ITDs and ILDs (i.e., when it was HRIR-convolved) but the noise source generated different signals for the two ears (Experiment 3). Altogether, our results demonstrate that binaurally coherent noise source location cues are necessary (Experiment 2; Bronkhorst and Plomp, 1988; Dieudonné and Francart, 2018; Garadat et al., 2009) but not sufficient (Experiment 3) for binaural unmasking to occur and that binaural noise source coherence is also a requisite for binaural unmasking to occur. To our knowledge, this has not been previously demonstrated experimentally, but our results could be probably predicted by binaural models that include noise (source) coherence as a factor that determines binaural unmasking (e.g., Culling et al., 2004, 2005). In addition, our results are consistent with studies that demonstrated the importance of noise source coherence for other binaural benefits. Noise source coherence is important to observe spatial release from masking (Culling et al., 2004; Swaminathan et al., 2016), to reduce the detrimental effects of reverberation on speech intelligibility (Lavandier and Culling, 2008), or to produce binaural benefits on speech intelligibility with headphones (Bernstein et al., 2015). For instance, Bernstein et al. (2015) presented unprocessed target and masker speech in the left ear together with unprocessed or vocoded (without TFS cues) speech masker in the right ear. Relative to the monaural condition, NH listeners showed greater increase in speech intelligibility when unprocessed rather than vocoded masker was added to the right ear. This result appears to be consistent with ours insofar as presenting a vocoded masker in only one ear would reduce the masker coherence in the TFS and thus the binaural advantage. A difference between the present study and that from Bernstein et al. (2015) is that they did not simulate listening in the free field and thus did not explore the relative importance of location and source coherence on binaural unmasking.

Users of bilateral cochlear implants (CIs) or a hearing aid and a CI (bimodal users) usually show small or absent binaural unmasking (Gantz et al., 2002; Kraaijenga et al., 2018; Schafer et al., 2011; Schleich et al., 2004; Van Hoesel and Tyler, 2003; Van Hoesel et al., 2002). This may occur because CI processing strategies remove TFS cues or because CI stimulation is not synchronized between the ears, something that hinders the encoding of ITDs (see Litovsky et al., 2017 for a more extensive review). It may also occur, however, because the use of two independent devices distorts the interaural noise coherence. Here, the results of Experiment 4 demonstrate that the use of a unilateral hearing aid reduces binaural unmasking in listeners who have normal binaural unmasking without the hearing aid. Because our hearing aid had a linear gain of 5 dB and a change in the masker level of 5 dB has no consequences on binaural unmasking (Experiment 2), and because ITDs were not present in the masker because stimuli were not HRIR convolved, our results demonstrate that binaural unmasking was reduced because the hearing-aid signal processing reduced the noise coherence between both ears. Therefore, binaural coherence should be better preserved with the use of two rather than one hearing device (hearing aids or CIs), especially if they are designed with the same number of frequency channels and processing. Preserving the interaural coherence, however, is not so simple and not always possible even when the pair of devices work in the same manner (e.g., Brown et al., 2016).

In summary, we have shown that binaural unmasking tended to decrease when altering binaural noise source location cues (ITDs), and that the decrease was even greater when providing incoherent noise source signals to the two ears [by making  $N_L(t) \neq N_R(t)$  in Eq. 1] or when processing the noise alone through a unilateral hearing device [by making  $D_L(t) \neq D_R(t)$  in Eq. 1]. Altogether, our results show that source location cues (ITDs) are necessary but not sufficient for binaural unmasking to occur and suggest that any aspect or mechanism that reduces the binaural coherence of the noise at the eardrums will likely reduce binaural unmasking.

#### 4. Experiment 5: the role of the medial olivocochlear reflex in binaural unmasking

##### 4.1. Rationale

This experiment was aimed at investigating if the MOCR contributes to binaural unmasking. In the preceding experiments, the words employed to assess binaural unmasking started 50 ms after the noise onset. This delay was likely insufficient for the MOCR to be fully active at the word onset because it takes at least 300 ms for the noise to fully activate this reflex (Backus and Guinan, 2006). In the present experiment, SRTs were measured for the same conditions as in Experiment 3 but using a noise-word onset delay of 500 ms rather than 50 ms to ensure the MOCR was fully active at the word onset. If the MOCR contributes to binaural unmasking, the magnitude of unmasking should be greater for the longer than for the shorter noise-word onset delay, i.e., unmasking should be greater here than in Experiment 3. In addition, because activation of the contralateral MOCR can enhance auditory nerve responses to tones in noise both when the contralateral noise is correlated or uncorrelated with the masking noise (e.g., Fig. 4 in Kawase and Liberman, 1993), the MOCR should contribute to binaural unmasking both for binaurally coherent and incoherent noises.

##### 4.2. Methods

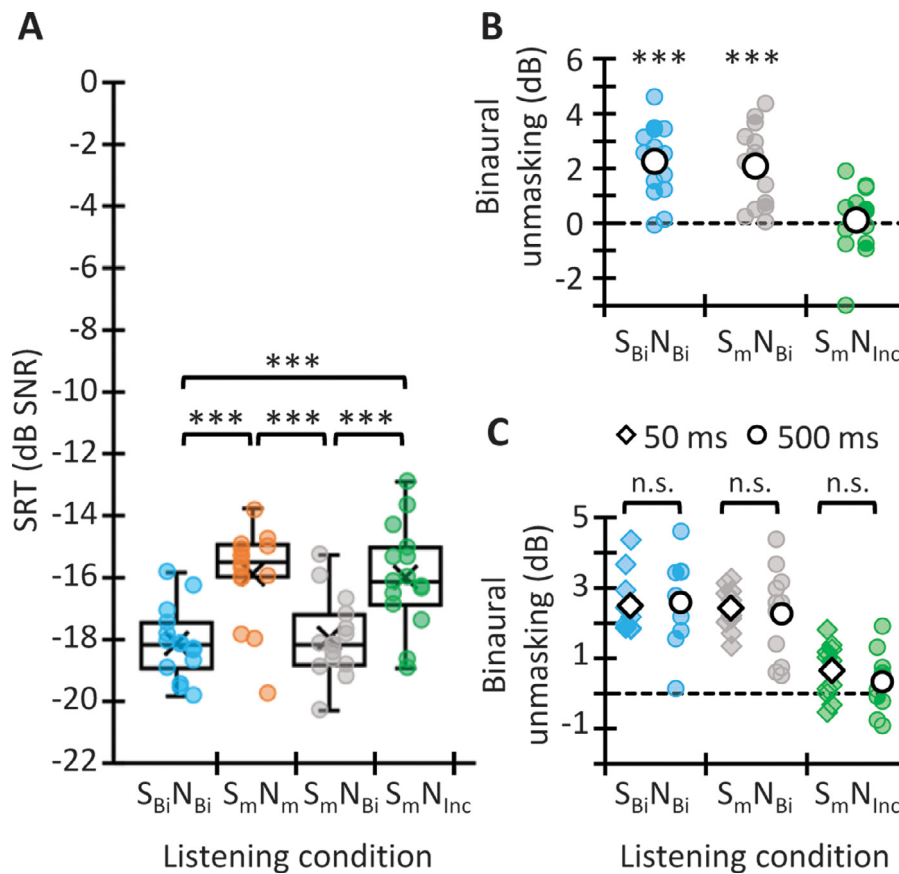
**Participants.** Fifteen NH listeners participated in this experiment (4 men; mean  $\pm$  SD age = 24.8  $\pm$  3.1 years). Eleven of them also participated in Experiment 3. Fourteen participants had audiometric thresholds  $\leq$  20 dB HL at both ears at octave frequencies between 250 Hz and 8 kHz. One participant had audiometric thresholds equal to 25 dB HL at 8 kHz in the right ear, and  $\leq$  20 dB HL at other frequencies. All of them were native Spanish speakers.

**Experimental conditions.** SRTs were measured for the same conditions as in Experiment 3, that is, listening with the two ears ( $S_{Bi}N_{Bi}$ ), listening with the acoustically better ear alone ( $S_mN_m$ ), listening with the two ears but removing the speech from the acoustically worse ear ( $S_mN_{Bi}$ ), and listening with the speech in the acoustically better ear and binaural incoherent noise ( $S_mN_{Inc}$ ). The difference is that in this experiment the noise started 500 ms instead of 50 ms before the words to allow full MOCR activation at the word onset. The noise finished 50 ms after the word offset.

**Stimuli, procedure, and apparatus** were as for Experiment 3. Binaural unmasking but not SRTs conformed to a normal distribution. A Friedman test was applied to test for the effect of listening condition on SRTs. A RMANOVA was used to test for the effect of listening condition on binaural unmasking. Post hoc pairwise comparisons were conducted using Bonferroni corrections for multiple comparisons. Two-tailed significance levels were applied for all tests.

##### 4.3. Results

Fig. 4A shows SRTs for the different conditions of Experiment 5. There was a significant effect of listening condition on SRT



**Fig. 4.** Results of Experiment 5. A. SRTs for 15 NH listeners for words delayed 500 ms from the noise onset. The layout is similar as Fig. 1A. B. Binaural unmasking calculated as the difference between SRTs in the monaural condition ( $S_mN_m$ ) and in the other test conditions of panel A. C. Binaural unmasking for words delayed 50 ms (Experiment 3; diamonds) or 500 ms (present experiment; circles) from the noise onset. Results are for 11 participants common to both experiments. Statistically significant pairwise comparisons are indicated by asterisks:  $***p \leq 0.001$ .

[ $N = 15$ ,  $\chi^2(3)=31.8$ ,  $p < 0.001$ ]. Participants showed binaural unmasking, i.e., better SRTs when they listened with their two ears than with their acoustically better ear alone ( $S_{Bi}N_{Bi}$  vs  $S_mN_m$  in Fig. 4A;  $\chi^2 = -2.0$ ,  $p < 0.001$ ). SRTs were not significantly different when listening to monaural speech in binaural noise compared to listening to binaural speech and noise ( $S_mN_{Bi}$  vs  $S_{Bi}N_{Bi}$  in Fig. 4A;  $\chi^2=0.0$ ,  $p = 1.000$ ). Binaural unmasking was absent when the noise was binaurally incoherent ( $S_mN_m$  vs  $S_mN_{Inc}$  in Fig. 4A;  $\chi^2=0.2$ ,  $p = 0.671$ ). Mean binaural unmasking was 2.3, 2.1 and 0.1 dB for  $S_{Bi}N_{Bi}$ ,  $S_mN_{Bi}$ , and  $S_mN_{Inc}$ , respectively (Fig. 4B). Altogether, the pattern of results was consistent with the pattern from Experiment 3.

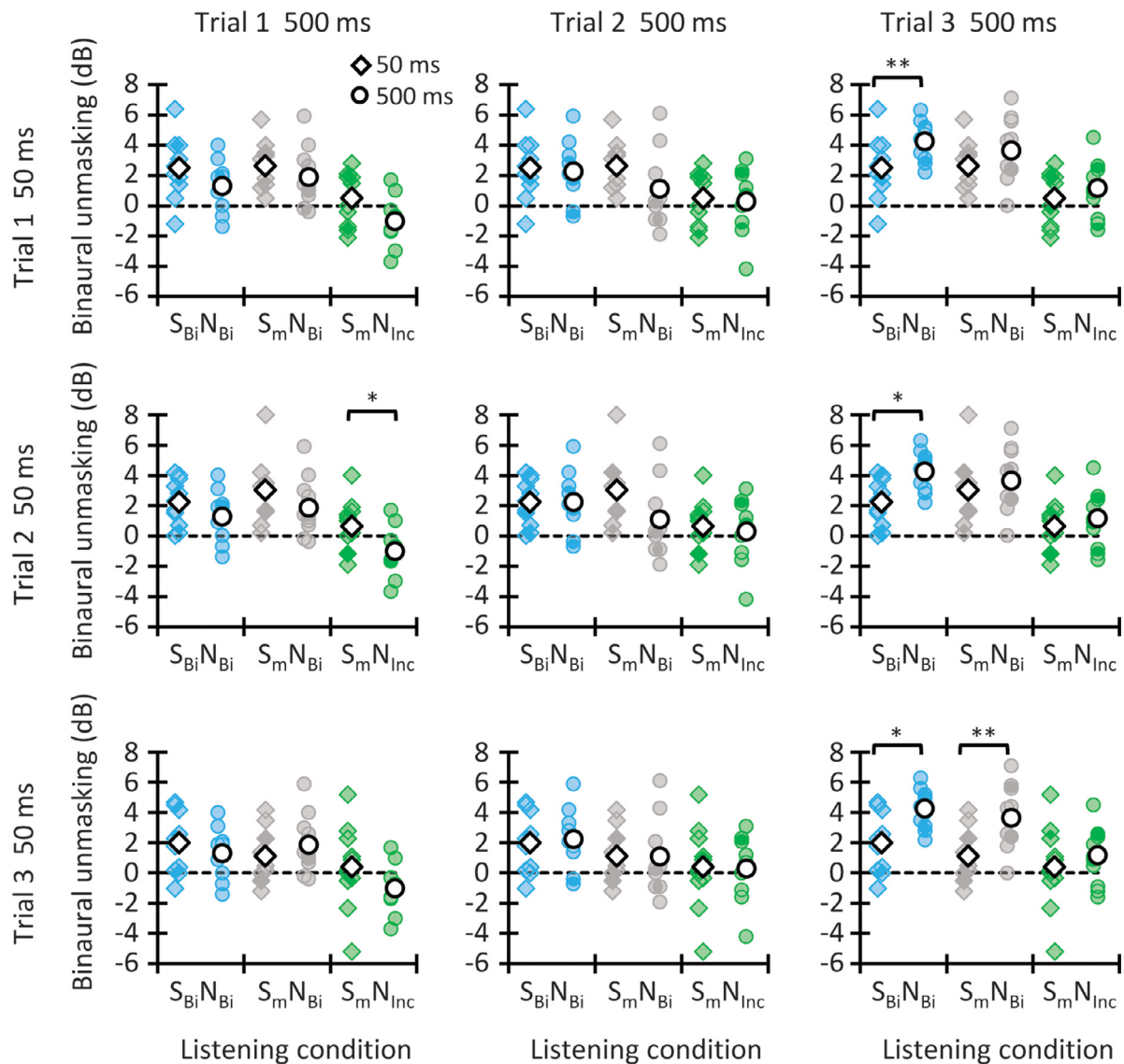
To compare binaural unmasking for noise-word onset delays of 50 ms (Experiment 3) and 500 ms (present experiment), we applied a RMANOVA with preceding noise duration (50 and 500 ms) and condition ( $S_{Bi}N_{Bi}$ ,  $S_mN_{Bi}$ , and  $S_mN_{Inc}$ ) as factors. Only data for the 11 participants that were common to the two experiments were included in the analysis (Fig. 4C). The RMANOVA revealed significant differences in the amount of binaural unmasking across condition [ $F(2,20)=41.4$ ,  $p < 0.001$ ]. However, there was not a significant effect of preceding noise duration [ $F(1,10)=0.2$ ,  $p = 0.675$ ] or an interaction between condition and noise duration [ $F(2,20)=0.4$ ,  $p = 0.649$ ] on binaural unmasking. The lack of an interaction indicates that binaural unmasking was not significantly different for words delayed 50 and 500 ms in any of the three conditions ( $S_{Bi}N_{Bi}$ ,  $S_mN_{Bi}$ , and  $S_mN_{Inc}$ ; Fig. 4C).

Recall that each individual SRT in Fig 4C (i.e., each filled symbol) is the mean of three SRT estimates, and so the findings just presented (about the lack of an effect of noise-word onset delay

on binaural unmasking) are based on mean SRTs. To assess the reliability of these findings, we repeated the analysis of the effect of word-onset delay on binaural unmasking for every possible pair of trials. Fig. 5 shows the amount of binaural unmasking for trials 1, 2, and 3 for words delayed 50 ms (diamonds) or 500 ms (circles) in the noise. Paired two-tailed  $t$ -tests without correction for multiple comparisons revealed that binaural unmasking was statistically different for words delayed 50 and 500 ms in the noise only for five out of the 27 pairwise comparisons (indicated by asterisks in Fig. 5), and the expected trends (i.e., greater unmasking for longer than for shorter noise-word onset delay) is observed in four of these comparisons. The greater binaural unmasking for the 500 ms condition in the third trial (right panels in Fig. 5), however, occurred because SRTs in the monaural condition worsened from trials one to three while SRTs in the binaural conditions remained constant across trials (data not shown). It is uncertain why this worsening occurred, but we expected the activation of ipsilateral and contralateral MOCR to improve SRTs in the binaural condition (re the monaural condition) rather than impair SRTs in the monaural condition.

#### 4.4. Discussion

The magnitude of binaural unmasking was not statistically different when words were delayed 50 and 500 ms from the noise onset (Fig. 4C). Assuming that the MOCR was minimally active at the word onset for words delayed 50 ms in the noise and fully active for words delayed 500 ms, the present results would indicate that the MOCR is unlikely involved in binaural unmasking. This as-



**Fig. 5.** Test-retest reliability of binaural unmasking. Binaural unmasking for trials 1, 2, and 3 when words were delayed 50 ms (diamonds) or 500 ms (circles) in the noise. Paired two-tailed *t*-tests without correction for multiple comparisons were used to test differences in binaural unmasking across the 50 and 500 ms conditions. Filled symbols depict individual results and open symbols depict group mean values. Statistically significant pairwise comparisons are indicated by asterisks: \**p* ≤ 0.05; \*\**p* ≤ 0.01. See the main text for details. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sumption, however, needs to be qualified. The evidence that the time course of activation of the MOCR is about 277±62 ms applies to probe tones ~1 kHz and wideband noise elicitors of the MOCR (Backus and Guinan, 2006) but the time course of activation might be different for other probe frequencies or elicitors. For instance, Maison et al. (2001) showed a MOCR time course of 60 ms for click probes in contralateral white noise and Kim et al. (2001) reported that the fast component of the ipsilateral MOCR can be highly variable across probe parameters and participants, with time courses ranging from 10 to 500 ms. If the MOCR were almost fully active at the word onset for some participants when words were delayed 50 ms in the noise, one would expect similar binaural unmasking for words delayed 50 or 500 ms, as we observed here. Note, however, that although this might have occurred, binaural unmasking was still approximately zero for binaurally incoherent noise. Because the antimasking effects of the MOCR on auditory nerve responses are similar for correlated and uncorrelated noise between ears (Kawase and Liberman, 1993), one would have expected binaural unmasking for incoherent noise if the MOCR were involved,

and this was not the case. Altogether, our results suggest that the MOCR is unlikely involved in binaural unmasking.

To our knowledge, this is the first study to directly investigate the potential contribution of the MOCR to binaural unmasking. Lopez-Poveda et al. (2020), however, have shown that binaural unmasking can occur for some simulated free-field conditions when stimuli are processed through a binaural CI sound-coding strategy inspired by the MOCR (Lopez-Poveda et al., 2016). Because CI users lack the effects and benefits of the natural MOCR, and this strategy is inspired by the natural MOCR and produces binaural unmasking, the findings of Lopez-Poveda et al. (2020) might be interpreted as indirect evidence that the MOCR contributes to binaural unmasking. The present findings do not support that idea. Instead, the present findings suggest that the better performance of CI users with the MOCR-inspired strategy is probably due to factors different from preserving MOCR-related effects (e.g., see the Discussion in Lopez-Poveda et al., 2020).

In addition to the MOCR, the middle ear muscle reflex (MEMR) might contribute to binaural unmasking. While the MOCR en-

hances the neural SNR at high frequencies, the MEMR possibly provides antimasking benefits in the low frequencies (Lieberman and Guinan, 1998). The MEMR threshold for ipsilateral or contralateral broadband noise activators can be as low as 50 dB SPL (Feeney et al., 2017; Zhao and Dhar, 2010). Here, SRTs were as low as -20 dB SNR for speech set at 65 dB SPL, indicating that the noise level was as high as 85 dB SPL and thus sufficient to activate the MEMR. The time course of the MEMR varies depending on the stimulus level and frequency (Borg, 1982; Qiu and Stucker, 1998), but it takes at least 240 ms for it to reach 90% of its maximum effect (Qiu and Stucker, 1998). Because binaural unmasking was similar for words delayed 50 or 500 ms in the noise, it is unlikely that the MEMR is involved in binaural unmasking.

## 5. General discussion

The present experiments have shown that: (1) in simulated free-field listening conditions, the noise in the ear with the worse acoustic SNR is necessary for binaural unmasking to occur, while the speech is dispensable (Experiment 1); (2) distortion of binaurally coherent noise source location cues (ITDs) tended to reduce binaural unmasking (Experiments 2); (3) the noise signals in the two ears must be binaurally coherent for binaural unmasking to occur (Experiments 3); (3) the use of a unilateral hearing aid with linear, minimal ( $\leq 5$  dB) amplification is sufficient to impair binaural unmasking for NH listeners, which suggests that the impairment is almost certainly due to the hearing aid distorting the interaural coherence of the noise stimuli at the eardrums (Experiment 4); and (4) the MOCR (and the MEMR) is unlikely to contribute to binaural unmasking (Experiment 5).

We found similar binaural unmasking in free-field listening conditions when listening to binaural speech in binaural noise as when listening to monaural speech in the better ear in binaural noise. In principle, this finding could be explained by two mechanisms at least. One possibility is that in free-field listening, the two ears function like a dual-microphone noise cancelling system (Fu et al., 2013). The acoustically better ear would operate like the primary microphone, which picks up the desired speech more strongly together with some background noise, while the worse ear would operate as a secondary microphone, which picks up mostly the noise. Just as a dual microphone system improves the SNR by subtracting the secondary from the primary signal, binaural unmasking could involve subtraction of the noise at the two ears so long as the noise is binaurally coherent (Durlach, 1963). The second mechanism would be related with spatial attention. Perhaps, with binaurally coherent noise, listeners were able to perceive a spatial separation between the target speech and noise source, which may have facilitated directing their attention to the acoustically better ear. If the listener attended mainly to speech presented in the better ear, intelligibility would be similar whether the speech was present or absent in the acoustically worse ear. According this second mechanism, binaural unmasking would be the result of listeners changing from not localizing with precision (monaural condition,  $N_m$ ) to correctly localizing the noise source (binaural condition,  $N_{Bi}$ ). The incoherent masker ( $N_{Inc}$ ), although HRIRs convolved, would have been more difficult to localize because it only preserved the onset ITDs and the long-term ILDs, thus resulting in less binaural unmasking.

The present results, however, suggest that noise cancellation is more prominently involved than spatial attention in binaural unmasking in free-field listening. This is because, for the  $S_mN_{Bi}$  condition, mean binaural unmasking was statistically significantly greater for stimuli not convolved than for stimuli convolved with HRIRs [3.2 vs 2.3 dB;  $t$ -test for independent samples;  $t(27)=2.6$ ,  $p = 0.016$ ] (compare the grey symbols in Fig. 3F and Fig. 3E). In other words, in the present experiments, unmasking was greater

for stimuli without than with source location information. This indicates that source location cues do not contribute to binaural unmasking, at least when speech is monaural and noise is binaural as in the present experiments. This result may seem inconsistent with studies that reported binaural unmasking to be greater for spatially separated than for collocated speech and masker sources (Bronkhorst and Plomp, 1988; Dieudonné and Francart, 2018; Garadat et al., 2009), which suggests that source location information is useful for binaural unmasking. In the latter studies, however, unmasking was always calculated as the difference between a bilateral condition ( $S_{Bi}N_{Bi}$ ) and the unilateral condition ( $S_mN_m$ ). They did not measure, however, binaural unmasking for separated and collocated speech and masker sources when speech is unilateral and noise is bilateral ( $S_mN_{Bi}$ ), which is precisely the condition that we tested and compared here. Whatever the mechanisms underlying binaural unmasking (noise cancellation or spatial attention), our findings show that the speech in the acoustically worse ear is dispensable for binaural unmasking to occur and thus dispute binaural speech intelligibility models that require the speech to be present in the acoustically worse ear (e.g., Tang et al., 2016). It is uncertain, however, if the same results would be obtained for all the possible locations of the speech and noise sources.

The use of a single hearing aid reduced binaural unmasking for NH listeners (Experiment 4). This suggests that hearing devices can reduce the binaural coherence of the noise even when their functioning is simple in terms of signal processing (our hearing aid had linear amplification rather than wide dynamic range compression that is commonly implemented in current hearing aids). Although previous studies suggest that the auditory brain can “learn” the coherence of the processed stimuli when the devices are used for a long period of time (Buss et al., 2008; Eapen et al., 2009; Litovsky et al., 2006; Mertens et al., 2015), the increase in binaural unmasking associated to this learning is small (0.5–1 dB in Litovsky et al., 2006; Mertens et al., 2015). Altogether, the present evidence shows that it would be unrealistic to expect normal binaural unmasking for users of one or two hearing devices. It further shows that it would be necessary that hearing devices preserve binaural (noise) coherence for their users to achieve closer to normal binaural unmasking (Litovsky et al., 2017).

As explained in the Introduction, based on the neural antimasking effects of the MOCR (e.g., Nieder and Nieder, 1970a, 1970b; Kawase and Liberman, 1993; reviewed by Lopez-Poveda, 2018) we expected the ipsilateral and contralateral MOCRs to improve the effective SNR in the acoustically better and worse ear, thus resulting in greater binaural unmasking when these reflexes are fully active than when they are not. We found, by contrast, that the amount of binaural unmasking was similar in the two cases. This suggests that the MOCR might not be as important for intelligibility in noise as it has been assumed and adds to recent evidence supporting the idea that physiological mechanisms not mediated by the MOCR facilitate speech-in-noise recognition (Marrufo-Pérez et al., 2018, 2020) or amplitude modulation detection in noise (Marrufo-Pérez et al., 2019; Wojtczak et al., 2019).

## Conclusions

- 1 In simulated free-field listening conditions, normal-hearing listeners show better intelligibility when listening with their two ears than when listening with their acoustically better ear alone, demonstrating binaural unmasking.
- 2 Having the noise in ear with the worse acoustic SNR is sufficient for binaural unmasking to occur, while the speech in the acoustically worse ear is dispensable. This suggests that binaural unmasking reflects binaural noise cancellation rather than speech enhancement in the central auditory system.

- 3 Interaural coherence pertaining to noise source location (ITDs) is necessary but not sufficient for binaural unmasking to occur. Interaural coherence is required also in the signals generated by the noise source and in their processing (e.g., by a hearing device).
- 4 The use of a unilateral hearing aid with linear, negligible (5 dB) amplification impaired binaural unmasking in normal-hearing listeners, almost certainly because its processing reduced the interaural coherence of the noise.
- 5 Binaural unmasking was not significantly different for stimuli designed to activate or not the medial olivocochlear reflex. This indicates that this reflex is unlikely involved in binaural unmasking.

### Declaration of Competing Interest

The authors report no conflict of interest.

### Acknowledgements

We thank Milagros J. Fumero for help with data collection, Thibaud Leclère for insightful discussions, and Peter T. Johannesen for help with hearing aid fitting. The hearing aid was kindly provided by Starkey Ltd (USA) in the context of a separate research project. Work supported by the Spanish Ministry of Economy and Competitiveness (grant PID2019-108985GB-I00) and by the European Regional Development Fund.

### Author statement

L.A.-S. and M.I.M.-P. performed research; L.A.-S. and M.I.M.-P. analysed data; A.E.-M. provided technical tools; E.A.L.-P. and M.I.M.-P. wrote the paper; M.I.M.-P. and E.A.L.-P. edited the paper; E.A.L.-P. designed research.

### References

Arsenault, M.D., Punch, J.L., 1999. Nonsense-syllable recognition in noise using monaural and binaural listening strategies. *J. Acoust. Soc. Am.* 105, 1821–1830.

American National Standards Institute, 1996. S3.6 Specification for Audiometers. American National Standards Institute, New York.

Backus, B.C., Guinan, J.J., 2006. Time-course of the human medial olivocochlear reflex. *J. Acoust. Soc. Am.* 119, 2889–2904.

Bernstein, J.W.G., Iyer, N., Brungart, D.S., 2015. Release from informational masking in a monaural competing-speech task with vocoded copies of the maskers presented contralaterally. *J. Acoust. Soc. Am.* 137, 702–713.

Borg, E., 1982. Time course of the human acoustic stapedius reflex: a comparison of eight different measures in normal-hearing subjects. *Scand. Audiol.* 11, 237–242.

Bronkhorst, A.W., Plomp, R., 1988. The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J. Acoust. Soc. Am.* 83, 1508–1516.

Brown, A.D., Rodriguez, F.A., Portnuff, C.D.F., Goupell, M.J., Tollin, D.J., 2016. Time-varying distortions of binaural information by bilateral hearing aids: effects of nonlinear frequency compression. *Trends Hear.* 20, 1–15.

Buss, E., Pillsbury, H.C., Buchman, C.A., Pillsbury, C.H., Clark, M.S., Haynes, D.S., Labadie, R.F., Amberg, S., Roland, P.S., Kruger, P., Novak, M.A., Wirth, J.A., Black, J.M., Peters, R., Lake, J., Wackym, P.A., Firszt, J.B., Wilson, B.S., Lawson, D.T., Schatzer, R., D'Haese, P.S., Barco, A.L., 2008. Multicenter U.S. bilateral MED-EL cochlear implantation study: speech perception over the first year of use. *Ear Hear.* 29, 20–32.

Cárdenas, M.R., Marrero, V., 1994. Cuaderno De Logaudiometría. Universidad Nacional de Educación a Distancia, Madrid, Spain.

Cooper, N.P., Guinan, J.J., 2006. Efferent-mediated control of basilar membrane motion. *J. Physiol.* 576, 49–54.

Culling, J.F., Hawley, M.L., Litovsky, R.Y., 2004. The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *J. Acoust. Soc. Am.* 116, 1057–1065.

Culling, J.F., Hawley, M.L., Litovsky, R.Y., 2005. Erratum: the role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *J. Acoust. Soc. Am.* 118, 552.

Culling, J.F., Summerfield, 1998. Measurements of the binaural temporal window using a detection task. *J. Acoust. Soc. Am.* 103, 3540–3553.

Dieudonné, B., Francart, T., 2018. Redundant information is sometimes more beneficial than spatial information to understand speech in noise. *Ear Hear.* 40, 545–554.

Durlach, N.I., 1963. Equalization and cancellation theory of binaural masking-level differences. *J. Acoust. Soc. Am.* 35, 1206–1218.

Eapen, R.J., Buss, E., Adunka, M.C., Pillsbury, H.C., Buchman, C.A., 2009. Hearing-in-noise benefits after bilateral simultaneous cochlear implantation continue to improve 4 years after implantation. *Otol. Neurotol.* 30, 153–159.

Feeney, M.P., Keefe, D.H., Hunter, L.L., Fitzpatrick, D.F., Garinis, A.C., Putterman, D.B., McMillan, G.P., 2017. Normative wideband reflectance, equivalent admittance at the tympanic membrane, and acoustic stapedius reflex threshold in adults. *Ear Hear.* 38 (3), e142–e160.

Fu, Z.H., Fan, F., Huang, J.D., 2013. Dual-microphone noise reduction for mobile phone application. IEEE International Conference on Acoustics, Speech and Target Processing doi:10.1109/ICASSP.2013.6639068.

Gantz, B., Tyler, R., Rubinstein, J., Wolaver, A., Lowder, M., Abbas, P., Brown, C., Hughes, M., Preece, J., 2002. Binaural cochlear implants placed during the same operation. *Otol. Neurotol.* 23, 169–180.

Garadat, S.N., Litovsky, R.Y., Yu, G., Zeng, F.G., 2009. Role of binaural hearing in speech intelligibility and spatial release from masking using vocoded speech. *J. Acoust. Soc. Am.* 126, 2522–2535.

Gardner, W.G., Martin, K.D., 1995. HRTF measurements of a KEMAR. *J. Acoust. Soc. Am.* 97, 3907–3908.

Guinan, J.J., 2006. Olivocochlear efferents: anatomy, physiology, function, and measurements of efferent effects in humans. *Ear Hear.* 27, 589–607.

Grantham, D.W., Whightman, F.L., 1979. Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. *J. Acoust. Soc. Am.* 65, 1509–1517.

Hauth, C.F., Brand, T., 2018. Modeling sluggishness in binaural unmasking of speech for maskers with time-varying interaural phase differences. *Trends Hear.* 22, 1–10. doi:10.1177/2331216517753547.

Johansson, M.S.K., Arlinger, S.D., 2002. Binaural masking level difference for speech targets in noise. *Int. J. Audiol.* 41, 279–284.

Kawase, T., Liberman, M.C., 1993. Antimasking effects of the olivocochlear reflex. I. Enhancement of compound action potentials to masked tones. *J. Neurophysiol.* 70, 2519–2532.

Kim, D.O., Dorn, P.A., Neely, S.T., Gorga, M.P., 2001. Adaptation of distortion product otoacoustic emission in humans. *J. Assoc. Res. Otolaryngol.* 2 (1), 31–40.

Kraaijenga, V.C.J., Tmit, A.L., Ramakers, G.G.J., Stokroos, R.J., van Zanten, G.A., 2018. No unmasking effect after sequential bilateral cochlear implantation in postlingually deafened adults: is there a first ear advantage? *Otol. Neurotol.* 39, e665–e670.

Koenig, W., 1950. Subjective effects in binaural hearing. *J. Acoust. Soc. Am.* 22, 61–62.

Kuhn, G.F., 1987. Physical Acoustics and Measurements Pertaining to Directional Hearing. Directional hearing. Springer, New York, NY, pp. 3–25.

Lavandier, L., Culling, J.F., 2008. Speech segregation in rooms: monaural, binaural, and interacting effects of reverberation on target and interferer. *J. Acoust. Soc. Am.* 123, 2237–2248.

Levitt, H., 1971. Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477.

Liberman, M.C., Guinan, J.J., 1998. Feedback control of the auditory periphery: anti-masking effects of middle ear muscles vs. olivocochlear efferents. *J. Commun. Disord.* 31, 471–483.

Licklider, J.C.R., 1948. The influence of interaural phase relations upon the masking of speech by white noise. *J. Acoust. Soc. Am.* 20, 150–159.

Litovsky, R.Y., Goupell, M.J., Misurelli, T.M., Kan, A., 2017. Hearing with cochlear implants and hearing aids in complex auditory scenes. In: Middlebrooks, J.C., Simon, J.Z., Popper, A.N., Fay, R.R. (Eds.), *The Auditory System At the Cocktail Party*. Springer Handbook of Auditory Research. Springer, New York, NY, pp. 261–291.

Litovsky, R.Y., Parkinson, A., Arcaroli, J., Sammath, C., 2006. Clinical study of simultaneous bilateral cochlear implantation in adults: a multicenter study. *Ear Hear.* 27, 714–731.

Lopez-Poveda, E.A., 2018. Olivocochlear efferents in animals and humans: from anatomy to clinical relevance. *Front. Neurol.* 9, 197.

Lopez-Poveda, E.A., Eustaquio-Martín, A., Stohl, J.S., Wolford, R.D., Schatzer, R., Wilson, B.S., 2016. A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Ear Hear.* 37, e138–e148.

Lopez-Poveda, E.A., Eustaquio-Martín, A., Fumero, M.J., Gorospe, J.M., Polo-López, R., Gutiérrez-Revilla, M.A., Schatzer, R., Nopp, P., Stohl, J.S., 2020. Speech-in-noise recognition with more realistic implementations of a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. *Ear Hear.* 41, 1492–1510.

Maison, S., Durrant, J., Gallineau, C., Micheyl, C., Collet, L., 2001. Delay and temporal integration in medial olivocochlear bundle activation in humans. *Ear Hear.* 22 (1), 65–74.

Marrufo-Pérez, M.I., Eustaquio-Martín, A., Lopez-Poveda, E.A., 2018. Adaptation to noise in human speech recognition unrelated to the medial olivocochlear reflex. *J. Neurosci.* 38, 4138–4145.

Marrufo-Pérez, M.I., Sturla-Carreto, D.P., Eustaquio-Martín, A., Lopez-Poveda, E.A., 2020. Adaptation to noise in human speech recognition depends on noise-level statistics and fast dynamic-range compression. *J. Neurosci.* 40, 6613–6623.

Marrufo-Pérez, M.I., Eustaquio-Martín, A., Fumero, M.J., Gorospe, J.M., Polo, R., Gutiérrez-Revilla, A., Lopez-Poveda, E.A., 2019. Adaptation to noise in amplitude modulation detection for cochlear implant users. *Hear. Res.* 377, 133–141.

Mertens, G., Kleine Punte, A., De Bodt, M., Van de Heyning, P., 2015. Binaural auditory outcomes in patients with postlingual profound unilateral hearing loss: 3 years after cochlear implantation. *Audiol. Neurootol.* 20, 67–72.

Murugasu, E., Russell, I.J., 1996. The effect of efferent stimulation on basilar mem-

- brane displacement in the basal turn of the guinea pig cochlea. *J. Neurosci.* 16, 325–332.
- Nieder, P., Nieder, I., 1970a. Stimulation of efferent olivocochlear bundle causes release from low level masking. *Nature* 227, 184–185.
- Nieder, P., Nieder, I., 1970b. Antimasking effect of crossed olivocochlear bundle stimulation with loud clicks in guinea pig. *Exp. Neurol.* 28, 179–188.
- Qiu, W.W., Stucker, F.J., 1998. Characterization of acoustic reflex latency in normal-hearing subjects. *Scand. Audiol.* 27, 43–49.
- Schafer, E.C., Amlani, A.M., Paiva, D., Nozari, L., Verret, S., 2011. A meta-analysis to compare speech recognition in noise with bilateral cochlear implants and bimodal stimulation. *Int. J. Audiol.* 50, 871–880.
- Schleich, P., Nopp, P., D'Haese, 2004. Head shadow, unmasking, and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant. *Ear Hear.* 25, 197–204.
- Swaminathan, J., Mason, C.R., Streeter, T.M., Best, V., Roverud, E., Kidd, G., 2016. Role of binaural temporal fine structure and envelope cues in cocktail-party listening. *J. Neurosci.* 36, 8250–8257.
- Tang, Y., Cooke, M., Fazenda, B.M., Cox, T.J., 2016. A metric for predicting binaural speech intelligibility in stationary noise and competing speech maskers. *J. Acoust. Soc. Am.* 140, 1858–1870.
- Van Hoesel, R., Tyler, R., 2003. Speech perception, localization, and lateralization with bilateral cochlear implants. *J. Acoust. Soc. Am.* 113, 1617–1630.
- Van Hoesel, R., Ramsden, R., O'Driscoll, M., 2002. Sound-direction identification, interaural time delay discrimination and speech intelligibility advantages in noise for a bilateral cochlear implant users. *Ear Hear.* 23, 137–149.
- Warr, W.B., Guinan, J.J., 1979. Efferent innervation in the organ of Corti: two separate systems. *Brain Res.* 173, 152–155.
- Winslow, R.L., Sachs, M.B., 1988. Single-tone intensity discrimination based on auditor-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. *Hear. Res.* 35, 165–190.
- Wojtczak, M., Klang, A.M., Torunsky, N.T., 2019. Exploring the role of medial olivocochlear efferents on the detection of amplitude modulation for tones presented in noise. *J. Assoc. Res. Otolaryngol.* 20, 395–413.
- Zhao, W., Dhar, S., 2010. The effect of contralateral acoustic stimulation on spontaneous otoacoustic emissions. *J. Assoc. Res. Otolaryngol.* 11, 53–67.