Speech-in-Noise Recognition With More Realistic Implementations of a Binaural Cochlear-Implant Sound Coding Strategy Inspired by the Medial Olivocochlear Reflex

Enrique A. Lopez-Poveda,[1,2,3] Almudena Eustaquio-Martín,[1,2] Milagros J. Fumero,[1,2]
José M. Gorospe,[1,2,3,4] Rubén Polo López,[5] M. Auxiliadora Gutiérrez Revilla,[5] Reinhold Schatzer,[6]
Peter Nopp,[6] and Joshua S. Stohl[7]

Objectives: Cochlear implant (CI) users continue to struggle understanding speech in noisy environments with current clinical devices. We have previously shown that this outcome can be improved by using binaural sound processors inspired by the medial olivocochlear (MOC) reflex, which involve dynamic (contralaterally controlled) rather than fixed compressive acoustic-to-electric maps. The present study aimed at investigating the potential additional benefits of using more realistic implementations of MOC processing.

Design: Eight users of bilateral CIs and two users of unilateral CIs participated in the study. Speech reception thresholds (SRTs) for sentences in competition with steady state noise were measured in unilateral and bilateral listening modes. Stimuli were processed through two independently functioning sound processors (one per ear) with fixed compression, the current clinical standard (STD); the originally proposed MOC strategy with fast contralateral control of compression (MOC1); a MOC strategy with slower control of compression (MOC2); and a slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels (MOC3). Performance with the four strategies was compared for multiple simulated spatial configurations of the speech and noise sources. Based on a previously published technical evaluation of these strategies, we hypothesized that SRTs would be overall better (lower) with the MOC3 strategy than with any of the other tested strategies. In addition, we hypothesized that the MOC3 strategy would be advantageous over the STD strategy in listening conditions and spatial configurations where the MOC1 strategy was not.

Results: In unilateral listening and when the implant ear had the worse acoustic signal-to-noise ratio, the mean SRT was 4 dB worse for the MOC1 than for the STD strategy (as expected), but it became equal or better for the MOC2 or MOC3 strategies than for the STD strategy. In bilateral listening, mean SRTs were 1.6 dB better for the MOC3 strategy than for the STD strategy across all spatial configurations tested, including a condition with speech and noise sources colocated at front where the MOC1 strategy was slightly disadvantageous relative to the STD strategy. All strategies produced significantly better SRTs for spatially separated than for colocated speech and noise sources. A statistically significant binaural advantage (i.e., better mean SRTs across spatial configurations and participants in bilateral than in unilateral listening) was found for the MOC2 and MOC3 strategies but not for the STD or MOC1 strategies.

Conclusions: Overall, performance was best with the MOC3 strategy, which maintained the benefits of the originally proposed MOC1 strategy over the STD strategy for spatially separated speech and noise sources and extended those benefits to additional spatial configurations. In addition, the MOC3 strategy provided a significant binaural advantage, which did not occur with the STD or the original MOC1 strategies.

Key words: Binaural advantage, Binaural hearing, Binaural sound processor, Olivocochlear efferents, Spatial masking release, Speech-in-noise intelligibility.

(Ear & Hearing 2020;41;1492–1510)

INTRODUCTION

Cochlear implants (CIs) are vastly successful but still open to improvement. Many users of CIs reach close-to-normal speech intelligibility in quiet environments (Wilson & Dorman 2007, 2008), but their intelligibility in noisy settings is still poorer than normal (Schleich et al. 2004; Loizou et al. 2009; Misurelli & Litovsky 2015; Wilson 2018). We have recently shown that for some listening conditions, the intelligibility of speech in competition with other sounds can be improved by using audio processors with binaurally coupled back-end compression inspired by the medial olivocochlear (MOC) reflex, an approach referred to as the "MOC strategy" (Lopez-Poveda et al. 2016a, 2017). Here, we report wider benefits of this strategy with more realistic implementations of the natural MOC reflex.

In healthy ears, the nonlinear mechanical vibration of the organ of Corti "maps" a wide range of acoustic pressure into a narrower (compressed) range of basilar membrane displacement (Robles & Ruggero 2001). The mapping, however, and thus the amount of compression, changes with activation of MOC efferents. MOC efferent activation suppresses the electromotility of outer hair cells in response to low-level sounds (Brown et al. 1983; Brown & Nuttall 1984). This linearizes basilar membrane input/output curves by inhibiting the amplitude of basilar membrane vibrations to low-level sounds without

<zdoi; 10.1097/AUD.0000000000000880>

significantly changing the response to high-level sounds (Murugasu & Russell 1996; Cooper & Guinan 2006). In quiet backgrounds, this linearization causes a mild increase in audiometric thresholds (Smith et al. 2000; Kawase et al. 2003; Aguilar et al. 2014). In noise, it restores the dynamic range of neural responses (Winslow & Sachs 1988) and releases neural responses from masking (Nieder & Nieder 1970), which presumably improves the neural coding of transient speech features and the intelligibility of speech in noise (see Lopez-Poveda 2018). Attention, as well as ipsilateral and contralateral sounds, can activate MOC efferents during natural listening, thereby adjusting compression dynamically and producing the "antimasking" effects just described. Normal-hearing individuals who have weak MOC reflexes have relatively poorer speech-in-noise perception (e.g., Mishra & Lutman 2014), which suggests that the antimasking effects of MOC reflex activation facilitate the intelligibility of speech in noise (see Lopez-Poveda 2018).

The electrical stimulation delivered by CIs is independent from MOC efferents, which might contribute to the greater difficulties experienced by CI users understanding speech in competition with other sounds compared with normal-hearing listeners. The MOC strategy was conceived to reinstate some efferent effects with CIs and other hearing devices (Lopez-Poveda 2015). Similar to the normal ear, the audio processor in a CI includes instantaneous compression at the back end in each frequency channel of processing to map a wide range of acoustic pressure into a narrower range of electrical current (Wilson et al. 1991, 2005; Wouters et al. 2015). The standard today is for this compression to be fixed (i.e., invariant over time). In the MOC strategy, by contrast, the amount of compression is conceived to change dynamically depending on control signals carefully selected to mimic attentional and/or reflexive efferent effects on compression (see Lopez-Poveda 2015; Lopez-Poveda et al. 2016b).

To date, the MOC strategy has been implemented and tested with contralateral control of compression to mimic the effects of the contralateral MOC reflex (attentional control and ipsilateral control of compression are foreseen but have not yet been investigated). The implementation involved on-frequency contralateral inhibition with short (2 msec) time constants for the activation and deactivation of the inhibition. Compared with using two independently functioning processors with fixed compression (the current clinical standard or STD), the MOC strategy enhanced the speech information in the ear with the better acoustic signal-to-noise ratio (SNR) (see later). As a result, the MOC strategy improved intelligibility for bilateral CI users when the target and interferer sound sources were spatially separated and for unilateral CI users when the implanted ear had the better acoustic SNR (see Lopez-Poveda et al. 2016a, 2017). The strategy, however, had potential drawbacks: (1) it reduced the speech information in the ear with the worse acoustic SNR, which could potentially hinder intelligibility in unilateral listening when the implant ear had the worse acoustic SNR (Note that the MOC strategy always involves two microphones (one per ear) and bilateral processing, as if users were wearing two CIs. In unilateral listening tests, the pattern of electrical stimulation is calculated for the two ears, but electrical stimulation is actually delivered only to the implant ear.); and (2) the mutual inhibition between the pair of processors decreased the overall stimulation levels and thus audibility, which could hinder intelligibility in bilateral or unilateral listening when the two CIs (or

processors) have identical input signals. (It is unlikely that bilateral CI users will have identical input signals at their implants in natural listening conditions. Identical inputs, however, can occur in well-controlled laboratory tests for colocated speech and interferer sources.)

The original implementation and parameters of the MOC strategy were chosen based on pilot comparisons of intelligibility for normal-hearing listeners presented with speech vocoded through the MOC and STD strategies (Lopez-Poveda & Eustaquio-Martin 2014). Such implementation and parameters disregarded aspects of the natural MOC reflex including the rather slow time courses for activation and deactivation of inhibition (Cooper & Guinan 2003; Backus & Guinan 2006), the possibility that the inhibition of basilar membrane responses be greater in apical than in basal cochlear regions (Lilaonitkul & Guinan 2009; Aguilar et al. 2013), and the possibility that the largest MOC reflex inhibition occurs when the contralateral sound elicitor is one-half octave below the probe frequency (Lilaonitkul & Guinan 2009). Lopez-Poveda and Eustaquio-Martín (2018) used the short-term objective intelligibility (STOI) to explore the potential benefits of MOC processing with more realistic implementations of natural MOC effects. STOI is an objective measure of the amount of information at the output of a sound processor (Taal et al. 2011). It is the average linear correlation (over time and frequency) between the unprocessed speech in quiet and the processed speech in noise. It is a scalar value between 0 and 1 that is expected to have a monotonic relation with the percentage of correctly understood speech tokens averaged across a group of listeners. The technical evaluation of Lopez-Poveda and Eustaquio-Martín predicted that the use of longer time constants for activation and deactivation of contralateral inhibition, combined with comparatively greater inhibition in the lower-frequency than in the higher-frequency channels, can overcome the shortcomings of the original MOC-strategy implementation and even improve the signal information in the ear with the worse acoustic SNR. In addition, the technical evaluation predicted no benefit of implementing a half-octave frequency offset in the contralateral control of inhibition.

The main aim of the present study was to experimentally confirm some of these predictions with actual CI users. A second aim was to investigate the potential binaural advantage provided by MOC processing. We measured speech reception thresholds (SRTs) for sentences presented in competition with steady state noise, in unilateral and bilateral listening modes, and for multiple spatial configurations of the speech and noise sources. SRTs were measured with the STD strategy, the "original" fast MOC strategy (MOC1), a slower MOC strategy (MOC2), and a slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels (MOC3). Measurements with a slower MOC strategy with offset contralateral control of inhibition were not conducted because of time constraints and because, as explained earlier, no benefits were expected from it. To verify the superior performance of the more realistic MOC implementations predicted by the STOI simulations of Lopez-Poveda and Eustaquio-Martín (2018), we included spatial configurations of the speech and noise sources where intelligibility was expected to be worse with the original MOC1 than with the STD strategy. All tests were conducted on eight bilateral and two unilateral CI users not previously tested on any of the strategies.

MATERIALS AND METHODS

The study was approved by the Ethics Review Board of the University of Salamanca.

Participants

Eight bilateral and two unilateral users of MED-EL CIs participated in the study (Table 1). Two of the bilateral CI users were children (SA012 and SA013), two were teenagers (SA009 and SA010), and four were adults (SA011, SA014, SA015, and SA016). The two unilateral CI users were adults (SA006 and SA007) and wore hearing aids in the ear contralateral to the CI. There was no particular reason for admitting participants of different ages to the study other than to increase the sample size (in Spain, adult bilateral CI users are scarce because the Spanish National Health Service covers bilateral implantation for children and only rarely for adults). This is unlikely problematic because all participants were able to perform the task and the study explored within-subject effects only (the main factors were processing strategy and spatial configuration). In other words, if any factor had made children perform differently from adults (e.g., Dubno et al. 2008; Eddins et al. 2018), the factor(s) in question would have affected all processing strategies equally.

All participants completed the whole set of tests except the two children and the unilateral CI users, who participated in a reduced number of conditions (see later). All participants were native speakers of Castilian Spanish. One of the children (SA013) had been living in Scotland for the last 4 years but he spoke Spanish at home. All participants were reported to perform very well with their implants. Participant SA009 had not been using his left implant for a month just before the start of the study because the audio processor was damaged.

Participants were volunteers and not paid for their service. They all signed an informed consent to participate in the study. None of them had been previously tested with any of the sound processing strategies used in the study.

Processing Strategies

Stimuli were processed through STD and MOC sound processing strategies before their presentation to participants. The STD and MOC strategies were identical to each other except for the back-end compression stage (Lopez-Poveda 2015; Lopez-Poveda et al. 2016a). The processors in the two strategies were based on the Continuous Interleaved Sampling (CIS) strategy (Wilson et al. 1991). They included a high-pass preemphasis filter (first-order Butterworth filter with a 3-dB cutoff frequency of 1.2 kHz); a bank of sixth-order Butterworth band-pass filters whose 3-dB cutoff frequencies followed a modified logarithmic distribution between 100 and 8500 Hz; envelope extraction via full-wave rectification and low-pass filtering (fourth-order Butterworth low-pass filter with a 3-dB cutoff frequency of 400 Hz); a logarithmic compression function (fixed for STD and dynamic for MOC processors); and CIS of the compressed envelopes with biphasic electrical pulses. The number of filters in the bank was identical to the minimum number of active electrodes between the left and right implants (Table 1) and equal for the left- and right-ear processors. The electrodes used for testing each participant are shown in Table 1.

The logarithmic compression function in all processors was as follows (Boyd 2006):

$$y = \frac{\ln(1 + c \cdot x)}{\ln(1 + c)}$$

(1)

TABLE 1. Participants' data

ID	Sex	Age (Years)	Etiology	Time of Implant Use (Months)		Electrodes Active Used		Pulse Rate (pps)		Better Ear	Thr (% MCL)	
				Left	Right	Left	Right	Left	Right		Left	Right
SA006	F	48	Genetic?	HA	125	n/a	1–11 1–11	n/a	1653	Right	n/a	5
SA007	M	49	Genetic?	HA	125	n/a	1–11 1–11	n/a	1617	Right	n/a	15
SA009	M	15	Genetic	105	148	1–12 1–10	3–12 3–12	1818	1538	Right	0	10
SA010	M	16	Unknown	140	172	1–12 1–10	1–10 1–10	1695	1099	Right	10	0
SA011	F	44	Antibiotic?	22	135	2–11 2–11	1–11 2–11	1754	1734	Left	5	5
SA012	F	7	Genetic	76	65	1–12 1–12	1–12 1–12	1515	1485	Left	5	5
SA013	M	8	Genetic	83	83	1–12 1–12	1–12 1–12	1485	1515	Right	10	10
SA014	M	48	Meningitis	175	190	1–9 1–9	1–7,9–11 1–7,9–10	1846	1143	Left	5	5
SA015	F	35	Meningitis	147	19	1–11 1–11	1–12 1–11	1405	1653	Left	5	5
SA016	F	74	Genetic?	150	119	1–10 1–10	1–2, 4–11 1–2, 4–11	1493	1478	Left	10	10

The better ear is as reported by the participant.
F, female; HA, hearing aid; M, male; MCL, maximum comfortable loudness; n/a, not applicable; pps, pulses per second; Thr, threshold.

where x and y are the input and output envelopes to/from the compressor, respectively, both assumed to be within the interval [0, 1]; and c is a parameter that determines the amount of compression.

STD Processors • For STD processors, c was set equal to 1000 and fixed. This value differed slightly from the value of 500 used by most of the participants in their clinical devices. The exceptions were the two unilateral CI users (SA006 and SA007), who were using $c = 1000$ in their clinical devices; the right-ear processor of SA010, which was configured with $c = 600$; the left-ear processor of SA014, which was configured with $c = 900$; and the left-ear processor of SA015, which was configured with $c = 1000$.

MOC Processors • In the MOC processors, the value of the compression parameter (c) in every frequency channel of processing varied dynamically depending upon the time-weighted output level from the corresponding frequency channel in the contralateral processor. The relationship between the instantaneous value of c and the instantaneous contralateral output level (E) was such that the greater the output level, the smaller the value of c (on-frequency inhibition). Specifically, c varied between approximately 30 and 1000 for contralateral output levels of 0 and −20 dB full scale (FS; where 0 dB FS means 0 dB re unity), respectively, as in the previously published experimental studies of the MOC strategy (Lopez-Poveda et al. 2016a, 2017).

Inspired by the exponential time course of activation and deactivation of the MOC reflex (Backus & Guinan 2006), in the MOC strategies, the instantaneous output level from the contralateral processor was calculated as the root-mean-square amplitude integrated over a preceding exponentially decaying time window with two time constants (τ_a and τ_b) (see later).

In previous experimental evaluations of the contralateral MOC strategy, the instantaneous compression parameter c for every frequency channel of processing depended upon the output level from the corresponding contralateral frequency channel (E). Due to the pseudologarithmic distribution of bandpass filter center frequencies, high-frequency channels had larger bandwidths than low-frequency channels. Therefore, for broadband signals, the output level and thus contralateral inhibition could have been greater for the higher-frequency than for the lower-frequency channels. To better control the amount of contralateral inhibition, after Lopez-Poveda and Eustaquio-Martín (2018), for the present MOC processors, the value of c for each frequency channel depended on the contralateral output level for the corresponding channel normalized to the channel bandwidth; that is, c depended on E' rather than E, where E' was calculated as follows:

$$E' = E \cdot \sqrt{\frac{BW_{ref}}{BW}}, \qquad (2)$$

where BW is the channel bandwidth and BW_{ref} is the bandwidth of a reference frequency channel.

Tested Strategies

SRTs were measured with the STD strategy and with three implementations of the MOC strategy. The latter involved dynamic and binaurally coupled back-end compression with different parameters:

- MOC1: This was the MOC strategy as implemented and tested originally (Lopez-Poveda et al. 2016b, 2017); that

is, with fast time constants ($\tau_a = \tau_b = 2$ msec) and with greater inhibition in the higher-frequency than in the lower-frequency channels (i.e., bandwidth normalization was not applied).
- MOC2: This was an MOC1 strategy with time constants $\tau_a = 2$ msec, $\tau_b = 300$ msec, thus overall closer to the slower time course of activation and deactivation of the natural contralateral MOC reflex (Backus & Guinan 2006).
- MOC3: This was an MOC2 strategy with bandwidth normalization to simulate greater inhibition in the apical than in the basal frequency channels, thus closer to the characteristics of the natural contralateral MOC reflex (Lilaonitkul & Guinan 2009). BW_{ref} was approximately equal to the bandwidth of median channel (the actual normalization channel was numbers 7, 6, 5, and 5 for participants with 12, 11, 10, and 9 active channels, respectively). As shown later, this produced effectively greater inhibition in the lower-frequency than in the higher-frequency channels.

Further details about these strategies can be found in Lopez-Poveda and Eustaquio-Martín (2018). The functioning of the various strategies is described later.

Equipment

The MATLAB software environment (R2014a; The Mathworks, Inc.) was used to perform all signal processing and implement all test procedures, including the presentation of electric stimuli. Stimuli were generated digitally (at 20 kHz sampling rate, 16-bit quantization), processed through the corresponding coding strategy, and the resulting electrical stimulation patterns delivered using the Research Interface Box 2 (Department of Ion Physics and Applied Physics at the University of Innsbruck, Innsbruck, Austria) and each patient's implanted receiver(s)/stimulator(s).

Speech Reception Thresholds

Intelligibility in noise was assessed by measuring the SNR at which listeners correctly recognized 50% of the full sentences that were presented. The resulting SNR will be referred to as the SRT. SRTs were measured using fixed-level speech (at −20 dB FS) and varying the noise level adaptively using a one-down, one-up procedure. For reference, the speech level of −20 dB FS corresponds approximately to 70 dB SPL in MED-EL clinical CI audio processors. For each SRT measurement, 30 sentences were presented and participants were asked to repeat each sentence. A sentence was scored as correct when all its words were correctly recognized and incorrect when at least one of the words was not recognized. The first 10 sentences were always the same but were presented in random order for all participants. They were included to give listeners an opportunity to become familiar with the processing strategy tested during the corresponding SRT measurement. The SNR changed in 3-dB steps for the first 14 sentences and in 2-dB steps for the final 17 sentences, and the SRT was calculated as the mean of the final 17 SNRs (the 31st SNR was calculated and used in the mean but not actually presented). If the SD of the 17 SNRs was greater than 3 dB, the SRT measurement was discarded and a new SRT was measured. Except for the two children (SA012 and SA013), three SRTs were measured in this way for each condition and the mean of the three measures was regarded as the final SRT. For the two children, only one SRT was measured per condition.

SRTs were measured using the Castilian Spanish version (Huarte 2008) of the hearing-in-noise test (HINT) (Nilsson et al. 1994) for a male target speaker. For the two children, SRTs were previously measured using the female sentences in the Spanish version of the Oldenburger Sentence Test (or "matrix" test) (Hochmuth et al. 2012). These SRTs, however, were regarded as part of the children's training in the SRT task and were discarded from further analyses. In all cases, the masker was speech-shaped HINT noise. A different noise token was used to mask each sentence. The noise started 500 msec before the sentence onset and ended 500 msec after the sentence offset and was gated with 50-msec cosine-squared onset and offset ramps.

Spatial Configurations

For unilateral CI users, SRTs were measured with the implanted ear alone (the hearing aid was removed during testing). For bilateral CI users, SRTs were measured in unilateral listening, involving listening with the self-reported better ear (Table 1), and in bilateral listening, involving listening with the two implants. SRTs were measured for five spatial configurations of the speech and noise sources in unilateral listening and for four spatial configurations in bilateral listening. Spatial configurations were different for different participants depending on the self-reported better ear of each participant. When the self-reported better was the right ear, unilateral listening was tested for S_0N_{60}, S_0N_0, S_0N_{-60}, $S_{15}N_{-15}$, $S_{60}N_{-60}$, and bilateral listening was tested for S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$, $S_{90}N_{-90}$. When the self-reported better ear was the left ear, unilateral listening was tested for S_0N_{-60}, S_0N_0, S_0N_{60}, $S_{-15}N_{15}$, $S_{-60}N_{60}$, and bilateral listening was tested for S_0N_0, $S_{-15}N_{15}$, $S_{-60}N_{60}$, $S_{-90}N_{90}$. In all cases, the speech and noise sources were at eye level (i.e., their elevation angle was 0°). In the S_XN_Y notation, X and Y indicate the azimuthal angles (in degrees) of the speech (S) and noise (N) sources, respectively, with 0° indicating a source directly in front and positive and negative values indicating sources to the right and the left of the midline, respectively. Note that locations were chosen so that the speech source was always in front or toward the self-reported better ear of each participant (i.e., spatial configurations were symmetrical about the midline for participants with different better ears). For convenience, in what follows, results are reported as if the better ear was the right ear for all participants.

Spatial locations were achieved by convolving monophonic recordings with diffuse-field equalized head-related transfer functions for a Knowles Electronics Manikin for Acoustic Research and for speakers 1 m away from the center of the manikin's head (Gardner & Martin 1995).

Order of Testing

Unilateral listening tests were always administered first followed by bilateral listening tests. For each of the two listening modes (bilateral or unilateral), measurements were organized in three blocks, one block for each of the three SRT estimates obtained per condition. In unilateral listening, each block involved measuring 20 SRTs (4 strategies × 5 spatial configurations). In bilateral listening, each block involved measuring 16 SRTs (4 strategies × 4 spatial configurations). Within each block, conditions were administered in random order, except for bilateral condition $S_{90}N_{-90}$, which was always administered last. Typically, a block was completed in two sessions separated by a short break. Sometimes, however, two or three sessions on consecutive days were needed to complete a block of measurements. If any individual SRT measurement did not meet the 3-dB SD criterion (see earlier), an additional SRT measurement was obtained after the full set of unilateral and bilateral tests was completed.

Neither the experimenter nor the participant knew of the strategy that was being tested at any time (double-blind approach).

The Castilian Spanish HINT corpus consists of 6 practice lists and 20 test lists with 10 sentences per list. Measuring each SRT required using one practice list plus two test lists. Therefore, the full protocol (adults and teenagers: 36 conditions × 3 SRT measurements per condition; children: 36 conditions × 1 SRT measurement per condition) involved using many more lists than were available. The lists used for each SRT measurement were selected randomly, but the procedure was designed so that all lists were used approximately the same number of times. The sentences in each list were presented in random order every time the list was used. The potential effects associated to reusing the lists are discussed later.

Fitting and Loudness Level Balance

Before testing, the electrical current levels at maximum comfortable loudness (MCL) were measured using the method of adjustment. Minimum stimulation levels (i.e., thresholds) were set to individually measured values or to 0%, 5%, or 10% of MCL values (Boyd 2006), according to each participant's preference (Table 1). Processor volumes were set using the STD strategy to ensure that sounds at the two ears were perceived as comfortable and equally loud and that a sentence filtered with the head-related transfer function for 0° elevation and 0° azimuth was perceived in the center of the head. A volume setting above 100% was required for some participants to achieve appropriate loudness levels. This resulted in a linear scaling up of the programmed levels for MCL in a fitting map. Threshold and MCL levels, as well as processor volumes, remained constant for each participant across conditions. They also remained constant for the MOC strategies to ensure that contralateral inhibition produced reductions in stimulation amplitudes (i.e., reduced loudness or audibility) relative to the STD strategy similar to those that the natural contralateral MOC reflex produces for listeners with normal hearing (Smith et al. 2000; Kawase et al. 2003; Aguilar et al. 2014).

Statistical Analyses

The results from unilateral and bilateral listening tests were analyzed separately. For each listening mode, a two-way repeated-measures analysis of the variance (RMANOVA) was conducted to test for the effects of processing strategy (STD, MOC1, MOC2, and MOC3), spatial configuration, and their interaction on group mean SRTs. The Greenhouse-Geisser correction was applied when the sphericity assumption was violated. Pairwise post hoc comparisons were conducted using Bonferroni corrections for multiple comparisons. All tests were two-tailed, and a result was regarded as statistically significant when $p \leq 0.05$. Statistical analyses were conducted using SPSS v. 23.
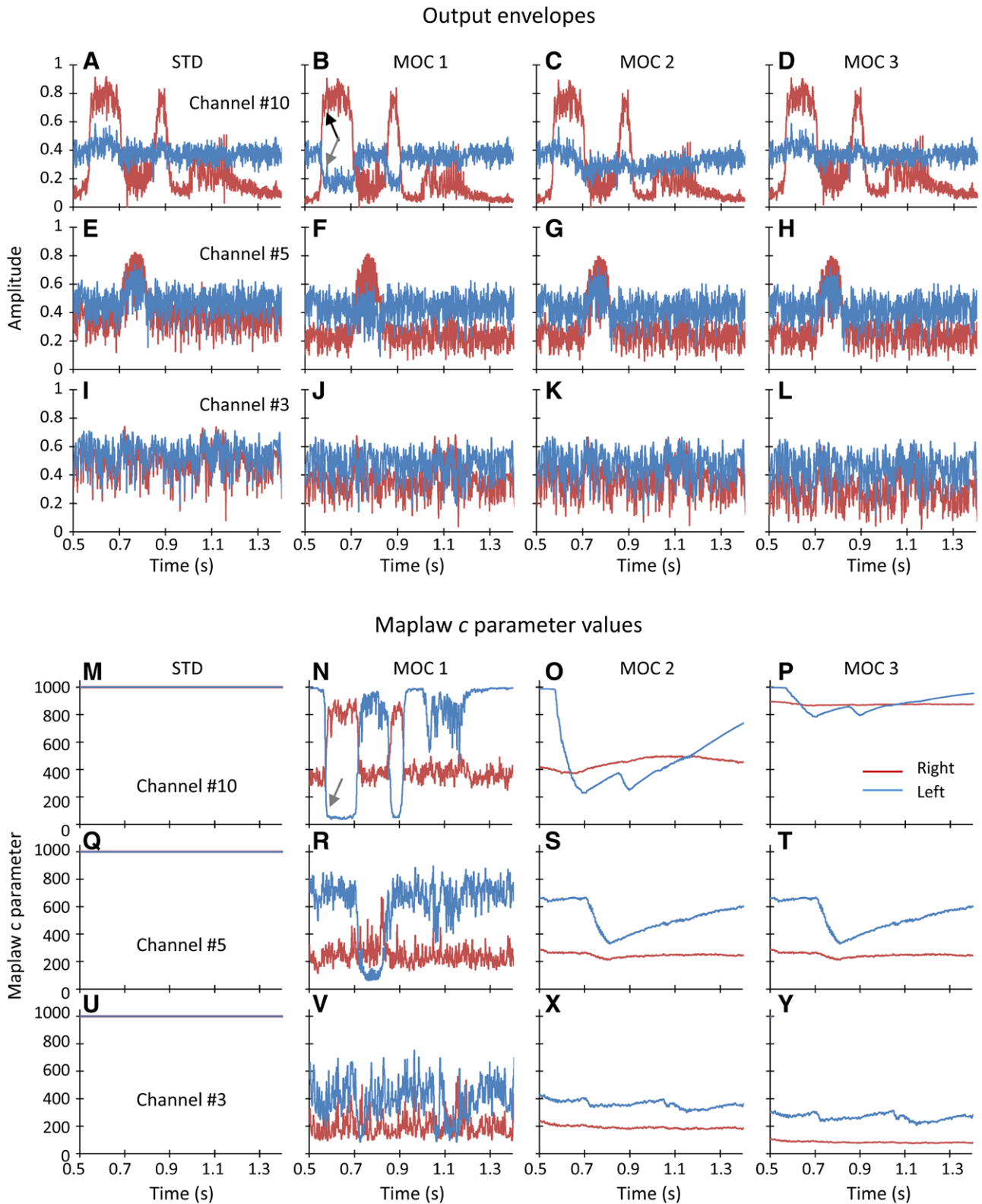
Fig. 1. Example compressed envelopes (A–L) and maplaw values (M–Y) for STD, MOC1, MOC2, and MOC3 strategies with 10 frequency channels. Data are shown only for three channels: channel number 3 (bottom row), channel number 5 (E–H and Q–T), and channel number 10 (top row) with center frequencies of 501, 1159, and 7230 Hz. The speech was the Castilian Spanish word "sastre," and the masker was speech-shaped noise. The speech and the masker had levels at −20 dB FS (i.e., 0 dB SNR) and were located at +60° and −60° azimuth, respectively. The masker started 500 msec before the speech. Red and blue traces show data for the right and left ears, respectively. Note the overlap between the red and blue traces in panels M, Q, and U, indicating that the value of the maplaw parameter c was equal across the ears in the STD strategy ($c = 1000$). FS indicates full scale; MOC, medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; SNR, signal-to-noise ratio; STD, standard.

Comparative Analysis of STD and MOC Output Envelopes

In this section, we illustrate the functioning of the tested strategies. The top part of Figure 1(panels A–L) shows output envelopes for STD, MOC1, MOC2, and MOC3 processors with 10 frequency channels, the typical number of channels used for the present participants (Table 1). For conciseness, output envelopes are shown only for three channels: channel numbers 3 (bottom row), 5 (middle row), and 10 (top row), with center frequencies of 501, 1159, and 7230 Hz, respectively. Blue and red traces illustrate envelopes for the left and the right ear, respectively. The speech was the Spanish word "sastre" and was located at +60° azimuth. The masker was speech-shaped noise and was located at −60° azimuth. The speech and noise had equal root-mean-square levels at −20 dB FS (i.e., 0 dB SNR) and the noise started 500 msec before the speech onset, as in the SRT measurements. The bottom part of Figure 1(panels M–Y) shows the corresponding time course of the maplaw (or compression) c parameter [Eq. (1)].

The figures show the following:

1. In the STD strategy, the maplaw parameter was constant ($c = 1000$), equal in the two ears, and equal across frequency channels. In the MOC1, MOC2, and MOC3 processors, by contrast, the maplaw parameter varied dynamically over time and was different across frequency channels and across ears.

2. The variation was such that when the amplitude in a given frequency channel was larger in one ear (black arrow in Fig. 1B), the maplaw c parameter and thus the amplitude decreased in the corresponding contralateral frequency channel relative to the STD strategy (gray arrows in Fig. 1B, 1N). In other words, the ear with the larger amplitude "inhibited" the ear with the smaller amplitude by decreasing the value of the maplaw parameter in the ear with the smaller amplitude.

3. The inhibitory effect, thus the temporal changes in the maplaw parameter, was faster for MOC1 than for MOC2 or MOC3 processors because the MOC1 strategy involved shorter (faster) time constants of contralateral inhibition than the MOC2 or MOC3 strategies.

4. For higher-frequency channels (channel number 10), which had larger bandwidths and thus produced higher output levels for broadband stimuli, inhibition was greater for MOC1 or MOC2 processors than for MOC3 processors (i.e., the maplaw parameter was overall smaller in Fig. 1N or Fig. 1O than in Fig. 1P). This is because unlike the MOC1 or MOC2 strategies, where parameter c depended on the raw contralateral output level, in the MOC3 strategy parameter, c depended on the contralateral output level normalized to the channel bandwidth [Eq. (2)].

5. For lower-frequency channels (channel number 3), inhibition was greater for MOC3 than for MOC2 processors (i.e., the maplaw parameter was slightly smaller in Fig. 1Y than in Fig. 1X) because of bandwidth normalization.

6. For the normalization frequency channel (channel number 5 in this example), the MOC2 and MOC3 processors had identical output envelopes (i.e., Fig. 1G was identical to Fig. 1H) and maplaw values (i.e., Fig. 1S was identical to Fig. 1T).

MOC processing can have several potential benefits over STD processing. To better understand some of those benefits, Figure 2 zooms in the output envelopes for channel number 5 (the channel best conveying the vowel /a/ in the word sastre) over the time period around the vowel /a/. Note that for this channel, MOC2 and MOC3 processors produced identical envelopes, hence the overlap between the green and purple traces. MOC processing involves greater contralateral inhibition for low than for high input levels (Lopez-Poveda et al. 2016a). In this example, the noise source was at −60° azimuth, hence closer to the left ear. Therefore, the higher noise levels in the left ear inhibited (reduced) the corresponding lower noise levels in right ear relative to the STD strategy at times before and after the vowel was present. Similarly, the higher vowel levels in the right ear inhibited (reduced) the corresponding vowel amplitudes in the left ear (recall that the speech source was at +60° azimuth, hence closer to the right ear). It is important to note that the reduction in vowel peaks was minimal in the ear closer to the speech source (the right ear). Altogether, this enhanced the effective SNR at the output of the MOC processors in the ear closer to the speech source, the right ear in this case (see also Fig. 3). In other words, the noise captured by the ear closer to the noise source (which had the worse acoustic SNR) contributed to enhancing the SNR in the ear closer to the speech source (which had the better acoustic SNR). That is, the acoustically worse ear made the acoustically better ear even better.

A second potential benefit from MOC processing is that it involves overall less compression, thus more linear processing than the STD processing (i.e., maplaw values are always equal or lower for the MOC than for the STD processors in Fig. 1). This is particularly true for the lower-frequency channels, where speech envelope cues are more salient. As shown by the inset in Figure 2A, this can enhance the representation of the vowel envelope, which is the acoustic cue that most current CI users rely on to understand speech.

The two benefits just described could be regarded as monaural benefits. A third potential benefit is binaural. The mutual inhibition involved in MOC processing can enhance the interaural level differences (ILDs) dynamically and on a channel-by-channel basis, as revealed by the fact that the maplaw values in Figure 1 were different for the two ears.

Figure 2 also serves to illustrate some of the main differences across MOC processors. Compared with an STD processor, MOC processing can reduce the speech level (thus the SNR) in the ear further away from the speech source. This is shown in Figure 2B, where the amplitudes over the time when the vowel was present were lower for the MOC1 strategy than for the STD strategy. This potentially detrimental effect, however, is less significant for the slower MOC2 or MOC3 processors than for the faster MOC1 processors (see also Fig. 3). In addition, the faster contralateral inhibition in the MOC1 strategy could potentially distort the speech envelopes more than the slower contralateral inhibition in the MOC2 or MOC3 strategies.

Figure 3 summarizes the effects and benefits of MOC processing just described by showing plots of compressed envelopes for different frequency channels as a function of time for the various processing strategies. Spatial color smoothing was used to improve the representation. The figure shows the following: (1) noise levels were overall lower for any MOC processor than for the STD processors, particularly in the right ear. (2) In the ear closer to the target source (the right ear in this
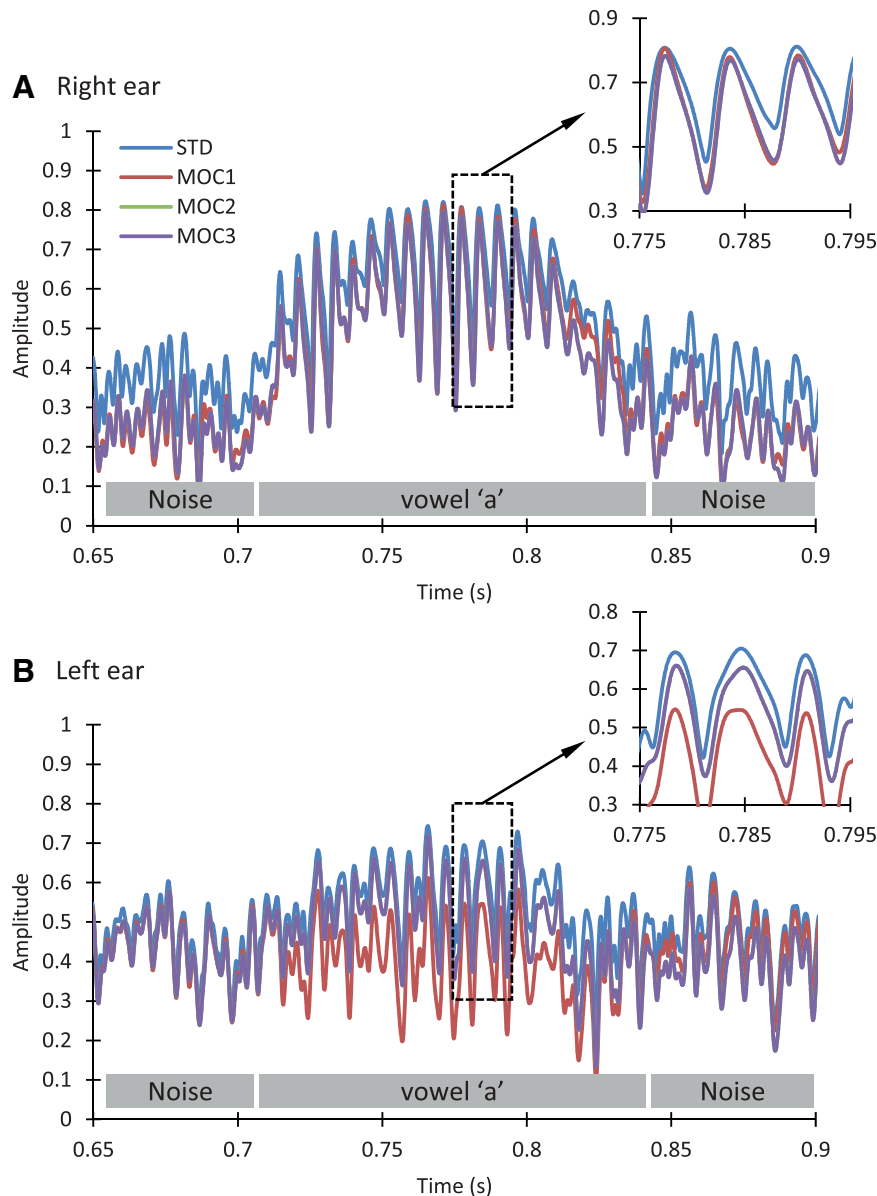
Fig. 2. Zoomed-in view of the compressed envelopes for channel number 5 shown in Fig. 1. Each panel shows envelopes for the STD, MOC1, MOC2, and MOC3 strategies. Envelopes were identical for the MOC2 and MOC3, hence the overlap between corresponding traces. The gray rectangles near the abscissae depict periods when the noise or the vowel /a/ were present. A, Envelopes for the right ear. B, Envelopes for the left ear. The inset in each panel illustrates a zoomed-in view of the envelopes over the area depicted by the corresponding rectangle. MOC indicates medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; STD, standard

example), the MOC strategies provided a better SNR than the STD strategy. (3) With MOC processing, some of the main speech features were inhibited in the left ear, particularly for the MOC1 and MOC2 strategies and less so for MOC3 strategy. As a result, the SNR in the left ear was higher for the MOC3 than for the MOC1 or MOC2 strategies. (4) In the right ear and in the lower-frequency channels (e.g., channel number 4), noise levels were lower for the MOC3 than for the MOC1, MOC2, or STD strategy. Altogether, it seems that the MOC3 processor provided the highest SNR in the right ear with minimal or no inhibition of speech cues in the left ear.

MOC processing can have one additional benefit (relative to STD processing) not seen in the output envelopes (not seen in Fig. 1, Fig. 2, or Fig. 3): the use of overall lower stimulation levels, particularly at times when noise was not present, could release auditory nerve neurons from adaptation, allowing them to better encode the speech envelope. Indeed, of the benefits just described, this neural antimasking effect is the main mechanism and benefit attributed to the MOC reflex in the literature (reviewed by Liberman & Guinan 1998; Lopez-Poveda 2018).

RESULTS

In this section, we first compare the SRTs for the various MOC strategies with those for the STD strategy in unilateral and bilateral listening. Then, we analyze the potential advantage
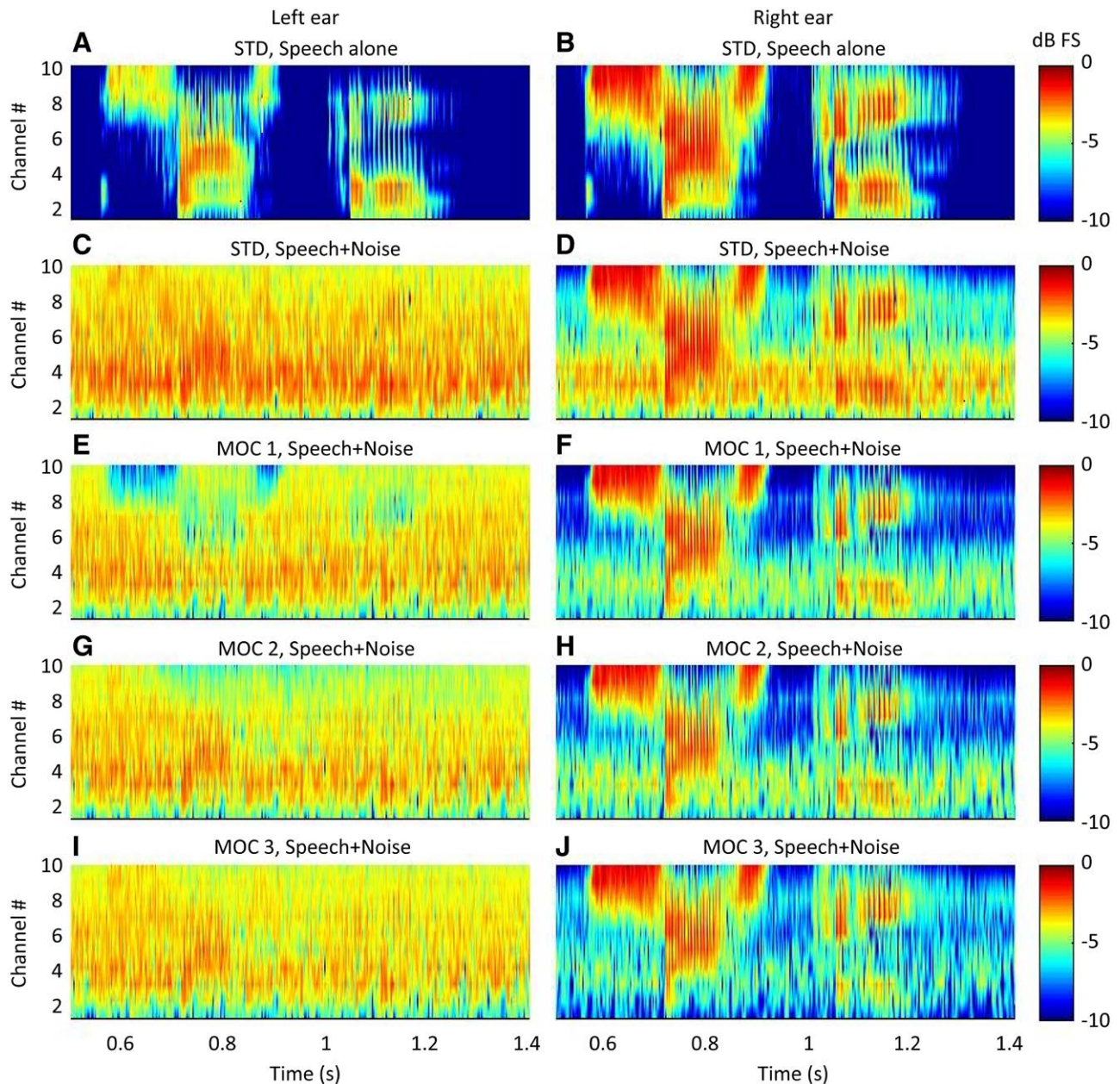
Fig. 3. Output envelopes for STD, MOC1, MOC2, and MOC3 processors with 10 frequency channels. The stimulus was as in Fig. 1. Each panel shows envelopes at the output of the maplaw as a function of frequency channel number and time. Color illustrates amplitude in units of dB FS, and spatial smoothing was applied to improve the view. Each row is for a different processing strategy, as indicated at the top of each panel. Left and right panels illustrate results for the left- and right-ear processors, respectively. As a reference, the top panels illustrate results for the STD strategy and for the word in quiet. All other panels illustrate results for the word and noise at −20 dB FS (0 dB SNR). FS indicates full scale; MOC, medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; SNR, signal-to-noise ratio; STD, standard.

of listening with two ears versus one ear with the tested processing strategies.

SRTs in Unilateral Listening

The top row in Figure 4 shows individual SRTs in unilateral listening (with the self-reported better ear) with the STD strategy. Each panel is for a different spatial configuration, as indicated at the top of each column. Recall that each value is the mean of at least three measurements, except for the two children

(SA012 and SA013) for whom only one SRT was obtained per spatial configuration. Rows 2 to 4 in Figure 4 illustrate the SRT improvement or "benefit" (in decibels) relative to the STD strategy provided by the MOC1, MOC2, and MOC3 strategies, respectively. The benefit was calculated as follows:

$$\text{SRT}_{\text{benefit}} [\text{dB}] = \text{SRT}_{\text{STD}} [\text{dB SNR}] - \text{SRT}_{\text{MOC}} [\text{dB SNR}] \quad (3)$$

Therefore, positive values indicate better intelligibility in noise (lower SRTs) with the corresponding MOC strategy than with

the STD strategy, while negative values indicate worse intelligibility (higher SRTs) with the MOC than with the STD strategy. Figure 5 shows group mean results.

For the S_0N_{60} spatial configuration (i.e., the most adverse listening condition with the speech source in front and the noise source at 60° toward the listening ear), the MOC1 strategy was disadvantageous for all participants (Fig. 4F). This is consistent with STOI simulations (see Fig. 5D in Lopez-Poveda & Eustaquio-Martín 2018) and was expected because the MOC1 strategy decreases the signal information in the ear contralateral to the speech source (compare the speech features in Fig. 3C and Fig. 3E). In contrast, SRTs were equal or better (up to 4 dB better for participant SA012) with the MOC2 than with the STD strategy (Fig. 4K) and equal or better (up to 2.3 dB better for participant SA015) with the MOC3 than with the STD strategy for all bilateral CI users (Fig. 4P). Even though the two unilateral CI users (SA006 and SA007, light color bars) did not benefit from MOC processing in this spatial configuration, their SRTs were nonetheless better with the MOC2 or MOC3 strategies than with the MOC1 strategy. On average, SRTs were 4.2 dB worse with the MOC1 than with the STD strategy but slightly better (<1 dB) with the MOC2 or MOC3 than with the STD strategy (Fig. 5B).

For speech and noise sources colocated in front of the participants (S_0N_0), many participants performed worse (up to 4.7 dB for participant SA009) with the MOC1 than with the STD strategy (Fig. 4G). This was expected based on earlier studies (Lopez-Poveda et al. 2016a) and STOI simulations (Fig. 5D in Lopez-Poveda & Eustaquio-Martín 2018) and possibly reflects reduced audibility and/or envelope distortion with the MOC1 strategy when the stimulus is identical at the two ears. By contrast, many participants benefited slightly from the MOC2 or the MOC3 strategies. Indeed, all bilateral CI users except SA012 showed equal or better SRTs with the MOC3 than with the STD strategy (Fig. 4Q). On average, SRTs were slightly worse with the MOC1 than with the STD strategy but slightly better with the MOC3 than with the STD strategy (Fig. 5B).

For the S_0N_{-60} spatial configuration (speech source in front with the noise source at 60° on the side contralateral to the CI), SRTs were generally worse with the MOC1 or MOC2 strategies than with the STD strategy (Fig. 4H, M). However, some participants benefited from the MOC3 strategy (Fig. 4R). This pattern of results was unexpected based on STOI simulations, which predicted SRT improvements of up to 6 dB for all MOC strategies (Fig. 5 in Lopez-Poveda & Eustaquio-Martín 2018). The reason for the discrepancy between the present experimental result and the STOI prediction is uncertain. STOI disregards the
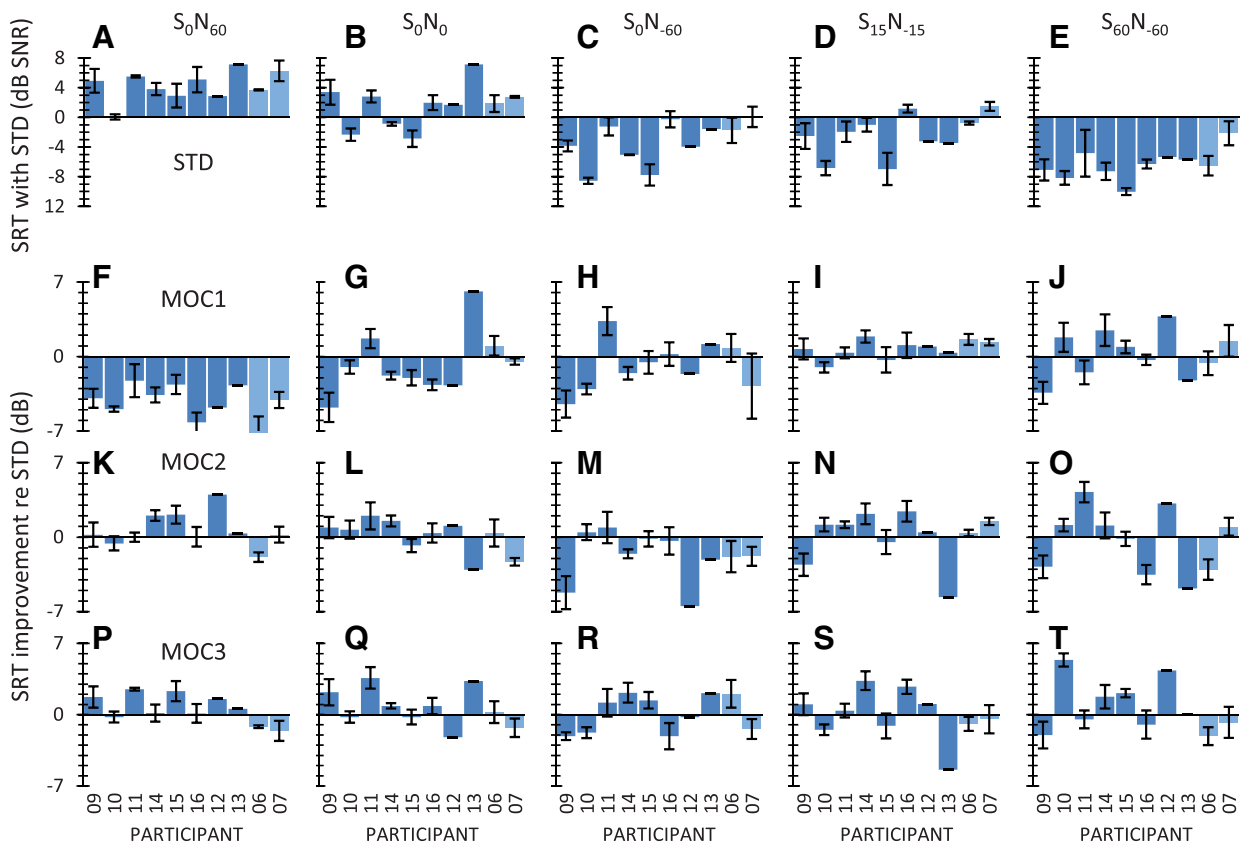

Fig. 4. Intelligibility in unilateral listening for individual participants. Row 1 (panels A to E), SRTs for the STD strategy. Each panel is for a different spatial configuration of the speech and noise sources, as indicated at the top. Rows 2 to 4 (panels F to T), SRT improvement relative to the STD strategy for the different MOC strategies (MOC1, MOC2, and MOC3). Data are shown for eight bilateral (darker bars) and two unilateral CI users (SA006 and SA007, lighter bars). Error bars illustrate 1 standard error of the mean. CI indicates cochlear implant; MOC, medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; N, noise; S, speech; SNR, signal-to-noise ratio; SRT, speech reception threshold; STD, standard.
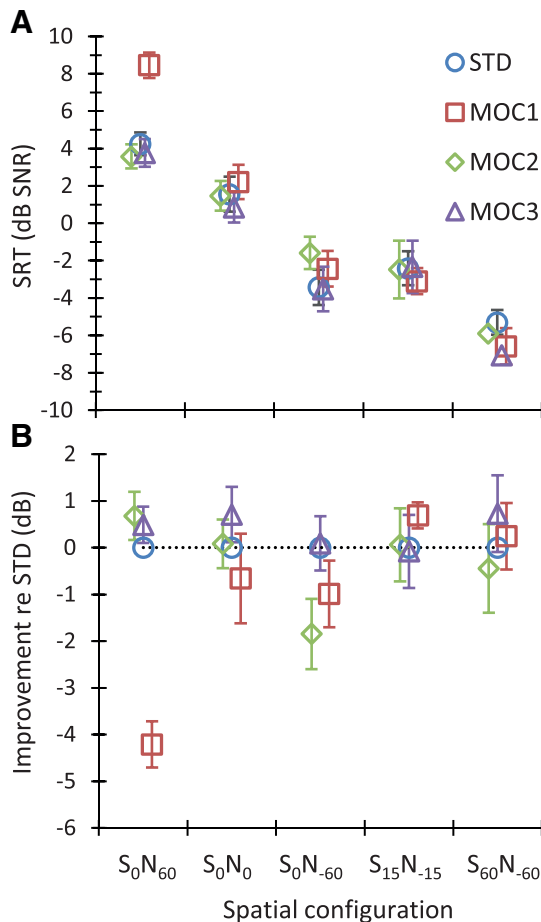
Fig. 5. Group mean intelligibility scores in unilateral listening. A, Mean SRTs for each strategy (as indicated by the inset) and spatial configuration (as indicated in the abscissa). Each point is the mean for eight bilateral and two unilateral CI users. B, Mean SRT improvement for the MOC strategies relative to the STD strategy. Error bars illustrate 1 standard error of the mean. CI indicates cochlear implant; MOC, medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; N, noise; S, speech; SRT, speech reception threshold; STD, standard.

effect of stimulation level on intelligibility, and the mutual inhibition between MOC processors causes stimulation level to be lower for the MOC than for the STD strategies. Therefore, perhaps, the speech level delivered by the MOC strategies was significantly more reduced in this than in other spatial configurations and hindered speech audibility.

For the $S_{15}N_{-15}$ and $S_{60}N_{-60}$ spatial configurations, some participants benefited from MOC processing, but others did not. Altogether, there was no clear benefit or disadvantage of MOC processing compared with STD processing (see also the mean SRT improvement in Fig. 5B).

A two-way RMANOVA was conducted to test for the effects of processing strategy (STD, MOC1, MOC2, and MOC3), spatial configuration (S_0N_{60}, S_0N_0, S_0N_{-60}, $S_{15}N_{-15}$, and $S_{60}N_{-60}$), and their interaction on the group mean SRTs. The RMANOVA revealed a significant effect of strategy [$F(3,27) = 4.34, p = 0.013$], spatial configuration [$F(2.5,22.1) = 190.60, p < 0.001$], and a significant interaction between processing strategy and spatial

configuration [$F(12,108) = 5.83, p < 0.001$]. A pairwise post hoc analysis with Bonferroni correction for multiple comparisons revealed that (1) the mean SRT for any strategy was not significantly different from the mean SRT for any other strategy ($p > 0.05$), except that the mean SRT was higher (worse) for the MOC1 than for the MOC3 strategies (−0.3 versus −1.7 dB SNR, $p = 0.027$); and (2) the mean SRT for any spatial configuration was different from the mean SRT for any other spatial configuration ($p \leq 0.001$), except S_0N_{-60} versus $S_{15}N_{-15}$ (mean SRTs across participants and processors were 5.0, 1.5, −2.7, −2.5, and −6.5 dB SNR for S_0N_{60}, S_0N_0, S_0N_{-60}, $S_{15}N_{-15}$, and $S_{60}N_{-60}$, respectively). Because SRTs tended to improve (become lower) with increasing the spatial separation between speech and noise sources, the latter confirmed that there was significant spatial release from masking.

A post hoc analysis of the interaction between strategy and spatial configuration showed a significant effect of processing strategy only for S_0N_{60} and produced the following p values: p(STD versus MOC1) < 0.001; p(STD versus MOC2) = 1.00; p(STD versus MOC3) = 1.00; p(MOC1 versus MOC2) < 0.001; p(MOC1 versus MOC3) < 0.001; and p(MOC2 versus MOC3) = 1.00. In other words, this analysis showed that for the S_0N_{60} spatial configuration (the most adverse listening condition with the speech source in front and the noise source at 60° toward the listening ear), the mean SRT was higher (worse) for the MOC1 strategy than for any other strategy (Fig. 5). For the other spatial configurations tested, the effect of strategy on SRT was not significant.

SRTs in Bilateral Listening

Figure 6 shows individual results in bilateral listening. The layout is the same as Figure 4. The top row shows individual SRTs for the STD strategy, while rows 2 to 4 illustrate the SRT improvement or benefit (in decibels) relative to the STD strategy provided by the MOC1, MOC2, and MOC3 strategies, respectively. Figure 7 shows corresponding group mean results.

For colocated speech and noise sources (S_0N_0 condition), the MOC1 strategy was disadvantageous compared to the STD strategy (the mean benefit was negative and equal to −0.9 dB, Figure 7), but the MOC2 and MOC3 strategies were beneficial (the mean SRT improvement was 1.7 and 1.8 dB, respectively). The MOC2 and MOC3 strategies were beneficial not only on average but also for most individual participants (Fig. 6I, M). The exception was SA010 with the MOC2 strategy. The benefit varied between 0 and 4 dB, depending on the participant. The largest benefits were for participant SA012 with the MOC2 and MOC3 strategies (3.9 and 4.0 dB, respectively).

For spatially separated speech and noise sources ($S_{15}N_{-15}$, $S_{60}N_{-60}$ and $S_{90}N_{-90}$ conditions), the group mean SRTs were better (lower) for all MOC strategies than for the STD strategy for all spatial configurations. With a few exceptions, a benefit was observed for each individual participant.

The RMANOVA test revealed a significant effect of strategy [$F(3,21) = 10.93, p < 0.001$] and spatial configuration [$F(1.43,10) = 87.27, p < 0.001$] on group mean SRTs. The interaction between strategy and spatial configuration was also significant [$F(9,63) = 2.83, p = 0.007$].

Post hoc pairwise comparisons, with Bonferroni correction, revealed that the SRTs measured with the MOC1, MOC2, and MOC3 strategies were not significantly different from each other
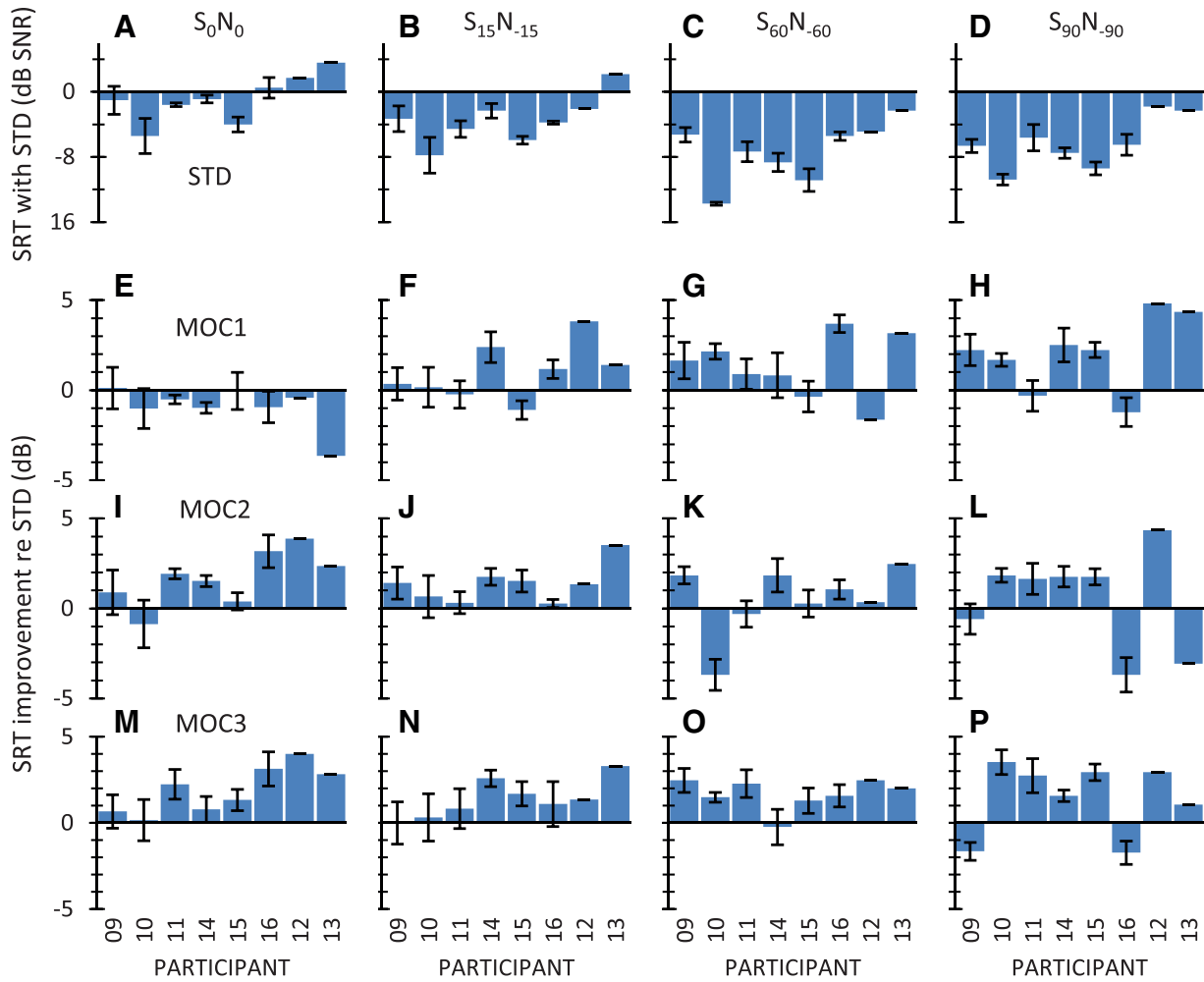
Fig. 6. Intelligibility in bilateral listening for individual participants. The layout is the same as Fig. 4. MOC indicates medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; N, noise; S, speech; SRT, speech reception threshold; STD, standard.

$[p(\text{MOC1 versus MOC2}) = 1.00; p(\text{MOC1 versus MOC3}) = 0.29;$ $p(\text{MOC2 versus MOC3}) = 0.50]$. In addition, it revealed that the SRTs for the MOC2 and STD strategies were not significantly different from each other $[p(\text{STD versus MOC2}) = 0.10]$. However, the mean SRT for the MOC1 strategy was significantly lower (better) than the mean SRT for the STD strategy (-5.3 versus -4.5 dB SNR, $p = 0.024$). The mean SRT for the MOC3 strategy was also significantly lower than the mean SRT for the STD strategy (-6.1 versus -4.5 dB SNR, $p = 0.003$). Indeed, except for the MOC1 at S_0N_0, the mean SRTs for all other conditions were lower (better) for the MOC1 and MOC3 than for STD strategy. This confirms that the MOC1 and MOC3 strategies produced significantly better speech-in-noise recognition than the STD strategy (Fig. 7).

Pairwise post hoc comparisons, using the Bonferroni correction, also revealed that SRTs were significantly different ($p < 0.05$) for every pair of spatial configurations except $S_{60}N_{-60}$ versus $S_{90}N_{-90}$ ($p = 0.10$). In other words, there was significant spatial release from masking between S_0N_0, $S_{15}N_{-15}$, and $S_{60}N_{-60}$, but not between $S_{60}N_{-60}$ and $S_{90}N_{-90}$.

Binaural Advantage

The term "binaural advantage" refers to the improvement in speech-in-noise intelligibility gained from listening with two ears compared with listening with one ear (e.g., Loizou et al. 2009; Avan et al. 2015). In this section, we address the question: what is the effect of the processing strategy on the binaural advantage?

The top panels in Figure 8 show the mean SRTs in noise for the STD, MOC1, MOC2, and MOC3 strategies in unilateral (open symbols) and bilateral listening (filled symbols) for the spatial configurations tested in the two listening modalities. Each data point is the group mean for the eight bilateral CI users. The bottom panels in Figure 8 show the difference between SRTs in unilateral minus bilateral listening (i.e., the binaural advantage). Overall, bilateral listening tended to be more advantageous over unilateral listening for spatially closer than for spatially separated speech and noise sources (recall that for spatially separated sources, the target was always closer to the self-reported better ear). For colocated speech and noise sources (S_0N_0 condition), bilateral listening tended to be more
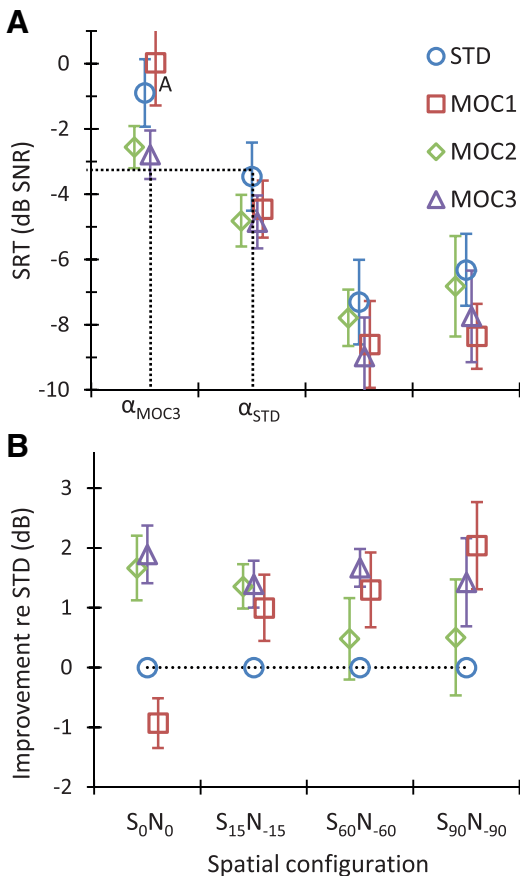
A

B

Fig. 7. Group mean intelligibility scores in bilateral listening. Each point is the mean for eight bilateral CI users. The layout is the same as Fig. 5. The dotted lines in panel A illustrate that at a fixed SNR of about −3 dB, the angular separation between the speech and noise source (α) to achieve 50% correct sentence recognition would be narrower for the MOC3 than for the STD strategy ($\alpha_{MOC3} < \alpha_{STD}$). CI indicates cochlear implant; MOC, medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; N, noise; S, speech; SNR, signal-to-noise ratio; SRT, speech reception threshold; STD, standard.

advantageous for the MOC2 and MOC3 strategies than for the MOC1 or STD strategies. For spatially separated speech and noise sources ($S_{15}N_{-15}$ and $S_{60}N_{-60}$ conditions), bilateral listening tended to be more advantageous for the MOC strategies than for the STD strategy.

An RMANOVA was conducted to test for the effects of listening modality (unilateral versus bilateral), spatial configuration (S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$), and their interaction on the group mean SRT. A separate test was conducted for each processing strategy. Table 2 shows the results. Significant effects are highlighted using bold font. SRTs decreased with increasing the spatial separation between the speech and noise sources, and the effect of spatial configuration was statistically significant for all four strategies. This shows that spatial release from masking was significant for all strategies. SRTs were equal or lower with two than with one CI, but the effect of listening modality was statistically significant only for the MOC2 and MOC3 strategies, indicating that only the MOC2 and MOC3 strategies provided a statistically significant binaural advantage.

The interaction between spatial configuration and listening condition was significant only for the MOC2 strategy, indicating that for this strategy, the binaural advantage depended on the spatial configuration. A post hoc comparison, using the Bonferroni correction method, indicated that for the MOC2 strategy, bilateral listening improved intelligibility when the speech and the noise sources were colocated (S_0N_0: $p = 0.007$) or separated by 30° ($S_{15}N_{-15}$: $p = 0.017$), but not when they were separated by 120° ($S_{60}N_{-60}$: $p = 0.210$).

Altogether, the present analysis demonstrates that only the MOC2 and MOC3 strategies produced a statistically significant binaural advantage, that is, better (lower) SRTs with two CIs than with one CI. The magnitude of the advantage decreased with increasing the spatial separation between the speech and noise sources.

A post hoc analysis of the data in Figure 8, with Bonferroni correction for multiple comparisons, revealed statistically lower (better) SRTs in bilateral than in unilateral listening for the S_0N_0 condition for the MOC2 ($p = 0.013$) and MOC3 ($p = 0.001$) strategies but not for the STD ($p = 0.061$) or the MOC1 ($p = 0.336$) strategy. In addition, it revealed better SRTs in bilateral than in unilateral listening for the $S_{15}N_{-15}$ condition for the MOC2 ($p = 0.031$) and the MOC3 ($p = 0.023$) strategies but not for the STD ($p = 0.975$) or the MOC1 ($p = 0.468$) strategies. For the $S_{60}N_{-60}$ condition, SRTs in bilateral listening were not statistically different from those in unilateral listening condition for any of the strategies (STD, $p = 0.829$; MOC1, $p = 0.437$; MOC2, $p = 0.534$; MOC3, $p = 0.354$). In other words, a binaural advantage was observed in the S_0N_0 and $S_{15}N_{-15}$ conditions but only with the MOC2 and MOC3 strategies and was not observed in the $S_{60}N_{-60}$ condition with any of the strategies.

DISCUSSION

We have shown in previous studies that, compared with using two independently functioning sound processors (STD strategy), the binaural MOC1 strategy improves SRTs for spatially separated speech and masker sources both in bilateral listening and in unilateral listening with the ear having the better SNR (Lopez-Poveda et al. 2016a, 2017). The MOC1 strategy, however, produces equal or worse SRTs for colocated speech and noise sources and theoretically can decrease the SNR in the ear with the worse acoustic SNR. The present study aimed at investigating if the benefits of MOC1 processing could be enhanced and its shortcomings overcome by using more realistic implementations of MOC processing, in particular, by using slower control of compression alone (MOC2 strategy) or combined with greater effects in the lower-frequency than in the higher-frequency channels (MOC3 strategy).

The main findings were as follows:

1. In bilateral listening and for spatially separated speech and noise sources, SRTs were better (lower) with the MOC1 than with the STD strategy (Fig. 7). This finding is consistent with the results of previous studies (Lopez-Poveda et al. 2016a, 2017).
2. In unilateral listening with the ear having the better SNR, SRTs were not significantly different for the MOC1 and the STD strategy for spatially separated speech and noise sources (Fig. 5). This may seem inconsistent with our previous study that reported the MOC1 to be advantageous over the STD strategy in similar conditions
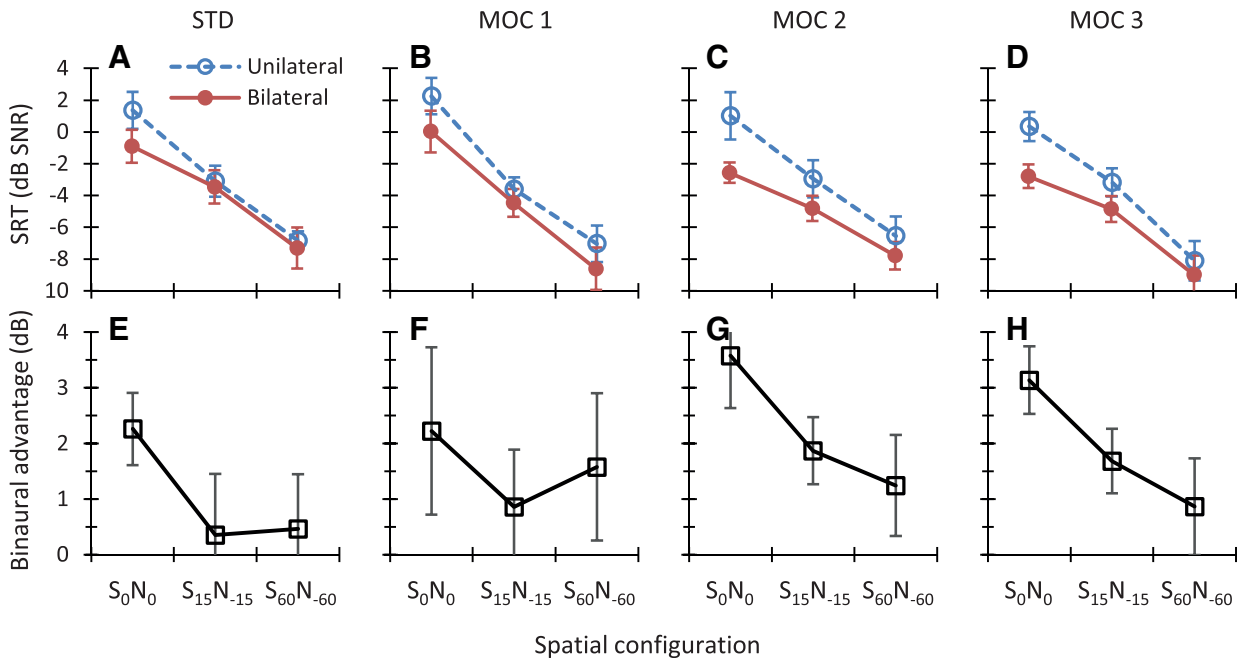
Fig. 8. Top, Group mean SRTs in unilateral and bilateral listening. Each panel is for a different strategy, as indicated at the top of the panel. Bottom, Mean binaural advantage calculated as the difference in mean SRT for unilateral listening minus bilateral listening. Positive values indicate better (lower) SRTs when listening with two rather one ear. Error bars illustrate 1 standard error of the mean. MOC indicates medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; N, noise; S, speech; SRT, speech reception threshold; STD, standard.

(Lopez-Poveda et al. 2016a). However, the spatial configurations were actually different for the two studies. Indeed, except for the S_0N_0 spatial configuration, none of the present unilateral listening conditions have been previously tested in combination with a speech-shaped noise masker.

3. In unilateral listening with the ear having the worse acoustic SNR (S_0N_{60} condition), SRTs were worse for the MOC1 than for the STD strategy but became equal or slightly better for the MOC2 or MOC3 strategies than for the STD strategy (Fig. 5). This finding confirms an expected, but yet untested, shortcoming of the MOC1 strategy (Lopez-Poveda et al. 2016a, 2016b). It also provides experimental support to a prediction made with STOI that the shortcoming in question can be overcome by using slower contralateral control of back-end compression (Lopez-Poveda and Eustaquio-Martín 2018).

4. In bilateral listening, the MOC1 strategy was advantageous over the STD strategy for spatially separated speech and noise sources but not for colocated speech and noise sources, where the mean SRT was slightly worse (0.9 dB higher) for the MOC1 than for the STD strategy (Fig. 7). The MOC3 strategy, however, was advantageous over the STD strategy for all spatial configurations tested, including the colocated condition. On average, the MOC3 strategy improved SRTs by 1.6 dB with respect to the STD strategy. This provides experimental support to a second prediction made with STOI that another shortcoming of the MOC1 strategy (namely, slightly worse SRTs relative to the STD strategy for colocated speech and noise sources) can be overcome by using slower control of compression combined with

greater effects in the lower-frequency than in the higher frequency channels (Lopez-Poveda and Eustaquio-Martín 2018).

5. All tested strategies (STD, MOC1, MOC2, and MOC3) produced significant spatial release from masking, both in unilateral (Fig. 5) and bilateral listening (Fig. 7) modes.

6. A statistically significant binaural advantage (i.e., better—lower—mean SRTs across spatial configurations and participants in bilateral than in unilateral listening) was found for the MOC2 and MOC3 strategies but not for the STD or MOC1 strategies (Fig. 8).

7. The binaural advantage with the MOC2 and MOC3 strategies was significant for colocated (S_0N_0) and spatially close ($S_{15}N_{-15}$) speech and noise sources but not for well-separated sources ($S_{60}N_{-60}$) (Fig. 8).

Compared with our earlier experimental studies of the MOC1 strategy, the present tests were conducted on a different group of CI users and involved additional spatial configurations of the speech and noise sources. Altogether the present data broadly confirm the benefits and shortcomings of the MOC1 strategy relative to STD strategy. They further show that the benefits of MOC1 processing may be enhanced and its shortcomings overcome by using more realistic implementations of MOC processing.

Spatial Release From Masking

Spatial release from masking (or the benefit obtained from separating the speech and noise sources in space) is often quantified as the difference in SRT for spatially colocated speech and noise sources (S_0N_0) minus the SRT for spatially separated

TABLE 2. Results of two-way RMANOVA tests for the effects of spatial configuration (S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$), listening modality (unilateral vs. bilateral listening), and their interaction on group mean SRTs

Strategy	N	Listening Modality	Spatial Configuration	Interaction
STD	8	$F(1,7) = 2.78, p = 0.139$	$\textbf{F(2,14) = 143.96, p < 0.001}$	$F(2,14) = 1.57, p = 0.240$
MOC1	8	$F(1,7) = 2.89, p = 0.130$	$\textbf{F(2,14) = 106.22, p < 0.001}$	$F(2,14) = 0.36, p = 0.700$
MOC2	8	$\textbf{F(1,7) = 10.36, p = 0.014}$	$\textbf{F(2,14) = 97.28, p < 0.001}$	$\textbf{F(2,14) = 4.32, p = 0.034}$
MOC3	8	$\textbf{F(1,7) = 20.22, p = 0.003}$	$\textbf{F(2,14) = 88.06, p < 0.001}$	$F(2,14) = 2.86, p = 0.091$

A separate test was conducted for each processing strategy (STD, MOC1, MOC2, and MOC3). Statistically significant effects are indicated using bold font.
MOC, medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; N, noise; RMANOVA, repeated-measures analysis of the variance; S, speech; SRT, speech reception threshold; STD, standard.

sources (see, for example, Fig. 4 in the review of Litovsky & Gordon 2016). According to this definition, the data in Figure 7 show that the mean spatial release from masking in bilateral listening for the $S_{60}N_{-60}$ versus S_0N_0 conditions was largest for the MOC1 strategy (8.6 dB), smallest for the MOC2 strategy (5.2 dB), and midrange and comparable for the STD (6.4 dB) and MOC3 (6.2 dB) strategies. Two comments are in order. First, spatial release from masking was largest for the MOC1 strategy because SRTs in the colocated condition were worst with this strategy. Second, the similarity between the magnitude of spatial release from masking for the STD and MOC3 strategies does not faithfully reflect the interaction between processing strategy and target-masker angular separation in situations where the SNR is fixed. Because mean SRTs for the reference condition (S_0N_0) were lower (better) for the MOC3 than for the STD strategy, at a fixed SNR, bilateral CI users would be able to recognize 50% of the sentences with a smaller angular separation when using the MOC3 than when using the STD strategy. For example, the dotted lines in Figure 7A illustrate that at −3 dB SNR, bilateral CI users would need speech and noise sources to be more widely separated with the STD than with the MOC3 strategy (approximately 30° versus 0°) to achieve 50% correct sentence recognition. Therefore, we would expect that in more realistic listening situations where the SNR and the speech-noise angular separations are both fixed, bilateral CI users would likely recognize a greater proportion of speech with the MOC3 than with the STD strategy.

Binaural Advantages of MOC Processing

Only the MOC2 and MOC3 strategies provided a statistically significant binaural advantage and only in the S_0N_0 and the $S_{15}N_{-15}$ conditions. A comparison of the present results with other studies (e.g., Tyler et al. 2002; Schleich et al. 2004; Litovsky et al. 2006; Buss et al. 2008; Loizou et al. 2009) is not straightforward because other studies involved different scoring (e.g., percent correct rather than SRT measurements), different spatial configurations (e.g., speech sources directly in front with noise sources on the sides), and/or users of clinical devices with several different technologies. Nonetheless, insofar as a comparison is possible, the present data for the STD strategy (the one closer to the current clinical standard in MED-EL devices) seem broadly consistent with those reported elsewhere. For example, Schleich et al. (2004) measured SRTs for 21 bilateral users of MED-EL clinical CIs in the free field and using the Oldenburg sentence test. For the S_0N_0 condition, they reported mean SRTs of −1.2 and 0.9 dB SNR in bilateral and unilateral listening, respectively, hence a binaural benefit of 2.1 dB. These values are not far from the present mean figures (SRTs of −0.9

and 1.4 dB SNR in bilateral and unilateral listening, respectively; and binaural benefit of 2.3 dB; Fig. 8E). In addition, for the S_0N_{-90} condition, Schleich et al. reported a mean SRT of −2.9 dB SNR when listening with the acoustically better ear (the right ear), which is not far from the mean SRT of −3.4 dB SNR for the most similar condition (unilateral listening in the S_0N_{-60} spatial configuration). Altogether, the similarity of the present data with the data of Schleich et al. supports the present findings and allows us to be optimistic that similar findings might be obtained in an eventual testing of the MOC strategies in the free field.

Compared with the STD strategy, the best MOC strategy (MOC3), and in general all MOC strategies, produced overall larger benefits in bilateral (Fig. 7) than in unilateral (Fig. 5) listening. The reason is unclear. The STD strategy was most similar to the audio processing strategies worn by the participants in their clinical devices, and unilateral listening tests were conducted before bilateral listening tests. Therefore, perhaps, participants were more used to MOC processing by the time that bilateral listening tests were conducted. This explanation, however, is not fully convincing because the pattern of results was broadly similar for the last block of unilateral listening tests (block number 3) and the first block of bilateral listening tests (block number 4), which were conducted consecutively. The pattern of results was also similar for the two last blocks of unilateral and bilateral listening tests (block numbers 3 and 6, respectively), when participants were presumably fully accustomed to the strategies.

An alternative interpretation for the greater benefit of MOC processing (relative to the STD strategy) in bilateral than in unilateral listening is that MOC processing provided little or no SNR improvement (relative to the STD strategy) in the ear with the better acoustic SNR but improved the SNR in the ear with the worse acoustic SNR and/or conveyed more natural binaural information. Of these two options, the first is unlikely to occur because, as shown in Figure 3 and by Lopez-Poveda and Eustaquio-Martín (2018), MOC processing reduces (MOC1) or slightly improves (MOC2 and MOC3) the speech information in the ear with the worse acoustic SNR. Indeed, when listening with the ear having the worse acoustic SNR (S_0N_{60} condition in Fig. 5), mean SRTs were worse for the MOC1 strategy or only slightly better for the MOC2 and MOC3 strategies than those for the STD strategy. Arsenault and Punch (1999) reported that normal-hearing listeners show better speech-in-noise recognition with natural binaural cues than when the stimulus at the ear with the better acoustic SNR is presented diotically. Therefore, the more parsimonious explanation for the greater benefit of MOC processing (relative to the STD strategy) in bilateral than in unilateral listening

is that MOC processing provided more natural binaural cues than the STD strategy.

Limitations

Given the limited number of sentence lists in the HINT corpus, we had to use the sentence lists multiple times to complete the comprehensive protocol. It is likely that participants learnt many of the sentences during testing. This may have turned the test from being "open set" at the beginning of testing to something more like "closed set" toward the end. As a result, the reported SRTs are probably lower than they would have been if we had not used the speech material repeatedly. We are confident, however, that reusing the sentences did not contribute to the reported differences in SRTs across strategies (or spatial configurations) because any one testing block involved testing all four processing strategies (and spatial configurations) in random order, before moving on to the next testing block. Therefore, the learning of the sentences and/or the improvement in performing the sentence recognition task would have affected all strategies and spatial configurations similarly.

The changing compression is central to MOC processing. It is known that different static compression values influence the SRT (e.g., Fu & Shannon 1998; Theelen-van den Hoek et al. 2016). Here, compression in the STD processor (i.e., the value of parameter c in Eq. (1)) was set to a (fixed) value that was not always the value used by the participants in their clinical processors (see Materials and Methods). Therefore, it remains unclear if any other static compression value would have resulted in better SRTs. In other words, one might wonder if the better performance with the MOC strategies may be due to a suboptimal STD compression setting. While possible, this is unlikely. First, we have previously shown that the MOC1 strategy can improve SRTs relative to the STD strategy both for steady state noise maskers (Lopez-Poveda et al. 2016a) and single-talker maskers (Lopez-Poveda et al. 2017), even when compression in the STD strategy is set equal to that used by the participants in their clinical audio processors. Second, we have previously shown that STOI scores, which are an objective, thus patient-independent measure of intelligibility, are greater with dynamic than with fixed compression, and STOI scores are well correlated with average patient performance (Lopez-Poveda and Eustaquio-Martín 2018). Third, Figure 9 shows that STOI scores (computed as described by Lopez-Poveda & Eustaquio-Martín 2018) are equal or higher for the MOC3 strategy than for an STD strategy set with $c = 500$, the typical value of the present participants in their clinical audio processors. Altogether, this suggests that the superior performance of MOC processing is unlikely due to a suboptimal compression setting in the STD strategy.

We note that the average benefits of MOC3 processing (Figs. 5, 7) held for many individual CI users (Figs. 4, 6). This seems remarkable considering that the STD strategy was the most similar to the audio processing strategies worn by the participants in their clinical devices and that participants were not given much opportunity to become fully accustomed to the MOC strategies before testing. For CI users, speech recognition can improve significantly over time and with training (e.g., Dorman & Spahr 2006) and some benefits of bilateral implantation

are seen only one year after the start of CI use (e.g., Buss et al. 2008). Therefore, it is tempting to speculate that the benefits from the MOC3 strategy could become larger with training and/ or a sustained use of the strategy.

Comparison With Other Binaural Algorithms and Final Remarks

There exist other sound processing approaches aimed at bringing the performance of bilateral CI users closer to that of listeners with normal hearing. Because the use of independent compression at the two ears can distort ILD cues and degrade speech-in-noise intelligibility (e.g., Wiggins & Seeber 2013), one approach consists of using linked (equal) automatic gain control (AGC) across the ears (e.g., Potts et al. 2019; Spencer et al. 2019). Compared with using unlinked AGC, the use of linked AGC can improve SRTs by 3.0 dB SNR for a speech source at 10° azimuth presented in competition with continuous four-talker babble at −70° azimuth (Potts et al. 2019). Another approach consists of preprocessing the acoustic stimuli binaurally before stimuli at the two ears are encoded into electrical pulses (reviewed by Baumgärtel et al. 2015a, 2015b). Binaural steering beamformers designed to track a moving sound source of interest in diffuse-field noise backgrounds can improve SRTs by about 4.5 dB (Adiloğlu et al. 2015), and other binaural preprocessing strategies can improve SRTs up to 10 dB when the target speech is presented in competition with single-talker maskers (reviewed by Baumgärtel et al 2015a, 2015b).

A direct comparison of the benefit provided by those approaches with that provided by MOC processing is hard because different studies have used different tasks, maskers, and/or spatial configurations. Insofar as a comparison is possible, however, the average SRT improvement provided by MOC processing (1.6 dB across the spatial configurations tested here) appears smaller than the benefit provided by those approaches. Binaural preprocessing strategies and beamformers, however, typically require the use of multiple microphones, speech detection and enhancement algorithms, and/or making assumptions about the characteristics of the target and/or the interferer sounds, or their spatial location (Baumgärtel et al. 2015b). By contrast, an implementation of the MOC strategy in a device would require one microphone per ear, no a priori assumptions about the signal of interest, no signal tracking, no complex preprocessing, and probably less exchange of data between the ears.

The MOC strategy can improve intelligibility over the STD strategy even when signals (and SNRs) are identical at the two ears, such as in the S_0N_0 condition (Fig. 7). This possibly reflects envelope enhancement due to the use of an overall more linear maplaw and/or neural antimasking associated to a reduced stimulation. Other benefits of MOC processing (see Materials and Methods), however, require an ILD, as provided by the head shadow. Insofar as the head-shadow ILDs can be reduced by the use of independent (unlinked) AGCs and natural ILDs may be somewhat restored by using linked AGC (Wiggins & Seeber 2013), MOC processing might provide larger benefits when used in combination with linked AGC. On the other hand, MOC processing, however, involves using dynamic rather than fixed acoustic-to-electric maps. The present evaluations involved implementing MOC processing in combination with a CIS sound coding strategy. As far as
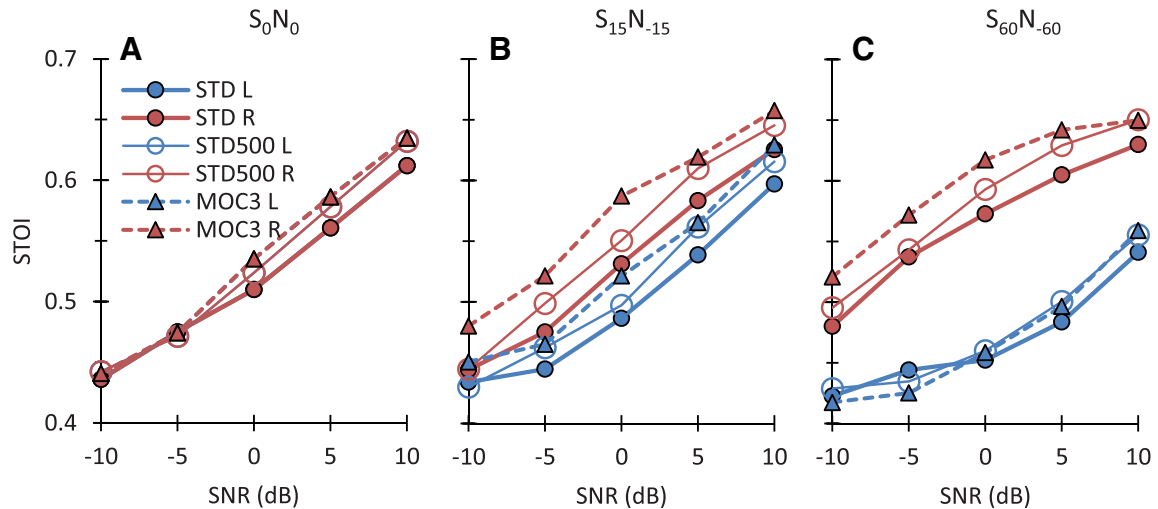
Fig. 9. Comparison of STOI scores for the present STD and MOC3 strategies against scores for a STD strategy with $c = 500$, the value typically used by the participants in their clinical audio processors. Each panel shows scores for the left (L) and right (R) ears (blue and red traces, respectively) and for different SNRs. Each panel is for a different spatial configuration of the target and speech sources, as indicated at the top of the panel. Note that for most SNRs and for the ear closer to the speech source (the right ear), STOI scores were equal or higher for the MOC3 strategy than for any of the two STD strategies. MOC indicates medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; N, noise; S, speech; SNR, signal-to-noise ratio; STD, standard; STOI, short-term objective intelligibility.

the authors know, however, all current sound coding strategies include acoustic-to-electric mapping at the back end of processing (see, for instance, Fig. 2 in Wouters et al. 2015). Therefore, MOC processing could be theoretically implemented with any CI sound coding strategy that does not already utilize dynamic back-end compression. Further research is necessary to investigate the potential benefits of combining MOC processing with linked AGC, with preprocessing beamformers, and with other sound coding strategies.

CONCLUSIONS

The SNR at 50% HINT sentence recognition was compared for CI users listening through experimental sound processing strategies involving the use of two independently functioning sound processors, each with fixed compressive acoustic-to-electric maps (the current clinical standard), or the use of binaurally coupled processors with contralaterally controlled dynamic compression inspired by the MOC reflex (the MOC strategy). Three versions of the MOC strategy were tested: an MOC1 strategy with fast contralateral control of compression (as proposed originally); an MOC2 strategy with slower control of compression; and an MOC3 strategy with slower control of compression and greater effects in the lower-frequency than in the higher-frequency channels. The main conclusions are as follows:

1. In unilateral listening, performance was worse with the MOC1 than with STD strategy when the listening ear had the worse acoustic SNR. By contrast, performance with the MOC2 or MOC3 strategies was comparable to that with the STD strategy in those same conditions.
2. In bilateral listening, performance was better with the MOC1 than with the STD strategy for spatially separated speech and noise sources but not for colocated sources.

The MOC3 strategy, however, was advantageous over the STD strategy for all spatial configurations tested, including the colocated condition. On average, the MOC3 strategy improved SRTs by 1.6 dB with respect to the STD strategy. This benefit was observed for most individual CI users.

3. The two main disadvantages of the MOC1 strategy relative to the STD strategy (namely, worse SRTs in bilateral listening for colocated speech and noise sources; and in unilateral listening when the listening ear had the worse acoustic SNR) were overcome by using longer time constants of activation and deactivation for the contralateral inhibition (i.e., with the MOC2 and MOC3 strategies).
4. All processing strategies produced significant spatial release from masking. However, in listening situations where the SNR and the angular separation between the speech and noise sources were both fixed, overall performance was best with the MOC3 strategy.
5. The MOC2 and MOC3 strategies produced a statistically significant binaural advantage, something that did not occur with the STD or MOC1 strategies.

Address for correspondence: Enrique A. Lopez-Poveda, Instituto de Neurociencias de Castilla y León, Calle Pintor Fernando Gallego 1, 37007 Salamanca, Spain. E-mail: ealopezpoveda@usal.es

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 Hear*, 19, 1–13.

Aguilar, E., Eustaquio-Martin, A., Lopez-Poveda, E. A. (2013). Contralateral efferent reflex effects on threshold and suprathreshold psychoacoustical tuning curves at low and high frequencies. *J Assoc Res Otolaryngol*, 14, 341–357.

Aguilar, E., Johannesen, P. T., Lopez-Poveda, E. A. (2014). Contralateral efferent suppression of human hearing sensitivity. *Front Syst Neurosci*, 8, 251.

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.

Avan, P., Giraudet, F., Büki, B. (2015). Importance of binaural hearing. *Audiol Neurootol*, 20(Suppl 1), 3–6.

Backus, B. C., & Guinan, J. J. Jr. (2006). Time-course of the human medial olivocochlear reflex. *J Acoust Soc Am*, 119(5 Pt 1), 2889–2904.

Baumgärtel, R. M., Hu, H., Krawczyk-Becker, M., Marquardt, D., Herzke, T., Coleman, G., Adiloğlu, K., Bomke, K., Plotz, K., Gerkmann, T., Doclo, S., Kollmeier, B., Hohmann, V., Dietz, M. (2015a). Comparing binaural pre-processing strategies II: Speech intelligibility of bilateral cochlear implant users. *Trends Hear*, 19, 1–18.

Baumgärtel, R. M., Krawczyk-Becker, M., Marquardt, D., Völker, C., Hu, H., Herzke, T., Coleman, G., Adiloğlu, K., Ernst, S. M., Gerkmann, T., Doclo, S., Kollmeier, B., Hohmann, V., Dietz, M. (2015b). Comparing binaural pre-processing strategies I: Instrumental evaluation. *Trends Hear*, 19, 1–16.

Boyd, P. J. (2006). Effects of programming threshold and maplaw settings on acoustic thresholds and speech discrimination with the MED-EL COMBI 40+ cochlear implant. *Ear Hear*, 27, 608–618.

Brown, M. C., & Nuttall, A. L. (1984). Efferent control of cochlear inner hair cell responses in the guinea pig. *J Physiol*, 354, 625–646.

Brown, M. C., Nuttall, A. L., Masta, R. I. (1983). Intracellular recordings from cochlear inner hair cells: Effects of stimulation of the crossed olivocochlear efferents. *Science*, 222, 69–72.

Buss, E., Pillsbury, H. C., Buchman, C. A., Pillsbury, C. H. Clark, M. S., Haynes, D. S., et al. (2008). Multicenter U.S. bilateral MED-EL cochlear implantation study: speech perception over the first year of use. *Ear Hear*, 29, 20–32.

Cooper, N. P., & Guinan, J. J. Jr. (2003). Separate mechanical processes underlie fast and slow effects of medial olivocochlear efferent activity. *J Physiol*, 548(Pt 1), 307–312.

Cooper, N. P., & Guinan, J. J. Jr. (2006). Efferent-mediated control of basilar membrane motion. *J Physiol*, 576(Pt 1), 49–54.

Dorman, M. F., & Spahr, A. J. (2006). Speech perception by adults with multichannel cochlear implants. In S. B. Waltzman & J. T. Roland (Eds.), *Cochlear Implants* (2nd ed.) (pp. 193–204). Thieme.

Dubno, J. R., Ahlstrom, J. B., Horwitz, A. R. (2008). Binaural advantage for younger and older adults with normal hearing. *J Speech Lang Hear Res*, 51, 539–556.

Eddins, A. C., Ozmeral, E. J., Eddins, D. A. (2018). How aging impacts the encoding of binaural cues and the perception of auditory space. *Hear Res*, 369, 79–89.

Fu, Q. J., & Shannon, R. V. (1998). Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners. *J Acoust Soc Am*, 104, 2570–2577.

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

Hochmuth, S., Brand, T., Zokoll, M. A., Castro, F. Z., Wardenaga, N., Kollmeier, B. (2012). A Spanish matrix sentence test for assessing speech reception thresholds in noise. *Int J Audiol*, 51, 536–544.

Huarte, A. (2008). The Castilian Spanish hearing in noise test. *Int J Audiol*, 47, 369–370.

Kawase, T., Ogura, M., Sato, T., Kobayashi, T., Suzuki, Y. (2003). Effects of contralateral noise on the measurement of auditory threshold. *Tohoku J Exp Med*, 200, 129–135.

Liberman, M. C., & Guinan, J. J. Jr. (1998). Feeback control of the auditory periphery: Anti-masking effects of the middle ear muscles vs. olivocochlear efferents. *J Commun Disord*, 31, 471–483.

Lilaonitkul, W., & Guinan, J. J. Jr. (2009). Human medial olivocochlear reflex: Effects as functions of contralateral, ipsilateral, and bilateral elicitor bandwidths. *J Assoc Res Otolaryngol*, 10, 459–470.

Litovsky, R. Y., & Gordon, K. (2016). Bilateral cochlear implants in children: Effects of auditory experience and deprivation on auditory perception. *Hear Res*, 338, 76–87.

Litovsky, R., Parkinson, A., Arcaroli, J., Sammath C. (2006). Simultaneous bilateral cochlear implantation in adults: a multicenter clinical study. *Ear Hear*, 27, 714–731.

Loizou, P. C., Hu, Y., Litovsky, R., Yu, G., Peters, R., Lake, J., Roland, P. (2009). Speech recognition by bilateral cochlear implant users in a cocktail-party setting. *J Acoust Soc Am*, 125, 372–383.

Lopez-Poveda, E. A. (2015). Sound enhancement for cochlear implants. EU Patent WO 2015/169649 A1.

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-Martin, A. (2014). Mimicking the unmasking effects of the medial olivo-cochlear efferent reflex with cochlear implants. *J Acoust Soc Am*, 135, 2385.

Lopez-Poveda, E. A., & Eustaquio-Martín, A. (2018). Objective speech transmission improvements with a binaural cochlear implant sound-coding strategy inspired by the contralateral medial olivocochlear reflex. *J Acoust Soc Am*, 143, 2217.

Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., Wilson, B. S. (2016a). 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., Stohl, J. S., Wolford, R. D., Schatzer, R., Wilson, B. S. (2016b). Roles of the contralateral efferent reflex in hearing demonstrated with cochlear implants. *Adv Exp Med Biol*, 894, 105–114.

Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., Gorospe, J. M., Ruiz, S. S. C., Benito, F., Wilson, B. S. (2017). Intelligibility in speech maskers with a binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Hear Res*, 348, 134–137.

Mishra, S. K., & Lutman, M. E. (2014). Top-down influences of the medial olivocochlear efferent system in speech perception in noise. *PLoS ONE*, 9, e85756.

Misurelli, S. M., & Litovsky, R. Y. (2015). Spatial release from masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity. *J Acoust Soc Am*, 138, 319–331.

Murugasu, E., & Russell, I. J. (1996). The effect of efferent stimulation on basilar membrane displacement in the basal turn of the guinea pig cochlea. *J Neurosci*, 16, 325–332.

Nieder, P., & Nieder, I. (1970). Stimulation of efferent olivocochlear bundle causes release from low level masking. *Nature*, 227, 184–185.

Nilsson, M., Soli, S. D., Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am*, 95, 1085–1099.

Potts, W. B., Ramanna, L., Perry, T., Long CJ. (2019). Improving Localization and Speech Reception in Noise for Bilateral Cochlear Implant Recipients. *Trends Hear*, 23, 2331216519831492.

Robles, L., & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiol Rev*, 81, 1305–1352.

Schleich, P., Nopp, P., D'Haese, P. (2004). Head shadow, squelch, and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant. *Ear Hear*, 25, 197–204.

Smith, D. W., Turner, D. A., Henson, M. M. (2000). Psychophysical correlates of contralateral efferent suppression. I. The role of the medial olivocochlear system in "central masking" in nonhuman primates. *J Acoust Soc Am*, 107, 933–941.

Spencer, N. J., Tillery, K. H., Brown, C. A. (2019). The effects of dynamic-range automatic gain control on sentence intelligibility with a speech masker in simulated cochlear implant listening. *Ear Hear, 40*, 710–724.

Taal, C. H., Hendriks, R. C., Heusdens, R., Jensen, J. (2011). An algorithm for intelligibility prediction of time-frequency weighted noisy speech. *IEEE Trans Audio Speech Lang Process, 19*, 2125–2136.

Theelen-van den Hoek, F. L., Boymans, M., van Dijk, B., Dreschler, W. A. (2016). Adjustments of the amplitude mapping function: Sensitivity of cochlear implant users and effects on subjective preference and speech recognition. *Int J Audiol, 55*, 674–687.

Tyler, R. S., Gantz, B. J., Rubinstein, J. T., Wilson, B. S., Parkinson, A. J., Wolaver, A., Preece, J. P., Witt, S., Lowder, M. W. (2002). Three month results with bilateral cochlear implants. *Ear Hear, 23*, 80S–89S.

Wiggins, I. M., & Seeber, B. U. (2013). Linking dynamic-range compression across the ears can improve speech intelligibility in spatially separated noise. *J Acoust Soc Am, 133*, 1004–1016.

Wilson, B. S. (2017). The cochlear implant and possibilities for narrowing the remaining gaps between prosthetic and normal hearing. *World J Otorhinolaryngol Head Neck Surg, 3*, 200–210.

Wilson, B. S., & Dorman, M. F. (2007). The surprising performance of present-day cochlear implants. *IEEE Trans Biomed Eng, 54*(6 Pt 1), 969–972.

Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hear Res, 242*, 3–21.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature, 352*, 236–238.

Wilson, B. S., Schatzer, R., Lopez-Poveda, E. A., Sun, X., Lawson, D. T., Wolford, R. D. (2005). Two new directions in speech processor design for cochlear implants. *Ear Hear, 26*(4 Suppl), 73S–81S.

Winslow, R. L., & Sachs, M. B. (1988). Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. *Hear Res, 35*, 165–189.

Wouters, J. W., McDermott, H. J., Francart, T. (2015). Sound coding in cochlear implants: From electric pulses to hearing. *IEEE Signal Process Mag, 2*, 67–80.