Hearing Research 409 (2021) 108320

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

A state-of-the-art implementation of a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex

Milagros J. Fumero[a,b], Almudena Eustaquio-Martín[a,b], José M. Gorospe[a,b,d], Rubén Polo López[e], M. Auxiliadora Gutiérrez Revilla[e], Luis Lassaletta[f,g,h], Reinhold Schatzer[i], Peter Nopp[i], Joshua S. Stohl[j], Enrique A. Lopez-Poveda[a,b,c,*]

[a] Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Calle Pintor Fernando Gallego 1, Salamanca 37007, Spain.
[b] Instituto de Investigación Biomédica de Salamanca, Universidad de Salamanca Salamanca 37007 Spain
[c] Departamento de Cirugía, Facultad de Medicina, Universidad de Salamanca, Salamanca 37007 Spain
[d] Servicio de Otorrinolaringología, Hospital Universitario de Salamanca, Salamanca 37007 Spain
[e] Servicio de Otorrinolaringología, Hospital Universitario Ramón y Cajal, Madrid 28034 Spain
[f] Servicio de Otorrinolaringología, Hospital Universitario La Paz, Madrid 28046 Spain
[g] IdiPAZ Research Institute, Madrid, Spain
[h] Biomedical Research Networking Centre on Rare Diseases (CIBERER-U761), Institute of Health Carlos III, Madrid, Spain
[i] MED-EL GmbH, Innsbruck, Austria
[j] North American Research Laboratory, MED-EL Corporation, Durham, NC, USA

ARTICLE INFO

Article history:
Received 16 April 2021
Revised 13 July 2021
Accepted 21 July 2021
Available online 24 July 2021

Keywords:
Binaural sound processing
Speech-in-noise intelligibility
Listening effort
Olivocochlear efferents

ABSTRACT

Cochlear implant (CI) users find it hard and effortful to understand speech in noise with current devices. Binaural CI sound processing inspired by the contralateral medial olivocochlear (MOC) reflex (an approach termed the 'MOC strategy') can improve speech-in-noise recognition for CI users. All reported evaluations of this strategy, however, disregarded automatic gain control (AGC) and fine-structure (FS) processing, two standard features in some current CI devices. To better assess the potential of implementing the MOC strategy in contemporary CIs, here, we compare intelligibility with and without MOC processing in combination with linked AGC and FS processing. Speech reception thresholds (SRTs) were compared for an FS and a MOC-FS strategy for sentences in steady and fluctuating noises, for various speech levels, in bilateral and unilateral listening modes, and for multiple spatial configurations of the speech and noise sources. Word recall scores and verbal response times in a word recognition test (two proxies for listening effort) were also compared for the two strategies in quiet and in steady noise at 5 dB signal-to-noise ratio (SNR) and the individual SRT. In steady noise, mean SRTs were always equal or better with the MOC-FS than with the standard FS strategy, both in bilateral (the mean and largest improvement across spatial configurations and speech levels were 0.8 and 2.2 dB, respectively) and unilateral listening (mean and largest improvement of 1.7 and 2.1 dB, respectively). In fluctuating noise and in bilateral listening, SRTs were equal for the two strategies. Word recall scores and verbal response times were not significantly affected by the test SNR or the processing strategy. Results show that MOC processing can be combined with linked AGC and FS processing. Compared to using FS processing alone, combined MOC-FS processing can improve speech intelligibility in noise without affecting word recall scores or verbal response times.

© 2021 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license
(http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

Cochlear implants (CIs) can enable useful hearing to deaf persons via direct electrical stimulation of the auditory nerve. Despite the progress achieved in CI design and performance, CI users still find it hard and effortful to understand speech in noise with modern bilateral CIs (BiCIs) (Schleich et al., 2004; Loizou et al., 2009; Hughes and Galvin, 2013; Wilson and Dorman, 2018; Hughes et al., 2018).

It has been shown that speech intelligibility in noise can be improved by providing CI users with a binaural CI sound-coding strategy termed 'the MOC strategy' (Lopez-Poveda et al., 2016a, 2016b).

---

[*] Corresponding author at: Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Calle Pintor Fernando Gallego 1 Salamanca 37007 Spain.
E-mail address: ealopezpoveda@usal.es (E.A. Lopez-Poveda).

https://doi.org/10.1016/j.heares.2021.108320
0378-5955/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

This strategy is inspired by (and named after) the dynamic control of basilar membrane compression provided in natural hearing by the contralateral medial olivocochlear reflex (MOCR). In contrast to the standard clinical approach (STD), which involves the use of two independently functioning audio processors with fixed acoustic-to-electric compression, the MOC strategy couples the two processors and dynamically adjusts the amount of compression applied in each ear. This can improve speech reception thresholds (SRTs) for speech in competition with steady noise (Lopez-Poveda et al., 2016b) and a single-talker masker (Lopez-Poveda et al., 2017) by more than 2 dB both in bilateral listening when the target and masker sources are spatially separated and in unilateral listening when the implanted ear has the better acoustic signal-to-noise ratio (SNR). In addition, it has been shown that speech-in-noise recognition is overall better with an implementation of the MOC strategy that reflects more realistically the characteristics of the natural MOCR (termed 'MOC3 strategy'), particularly a slower time course of contralateral inhibition and the greater inhibition in the lower frequency than in the higher frequency channels (Lopez-Poveda et al., 2020). This strategy maintained the benefits of the originally proposed MOC strategy over the STD strategy for spatially separated speech and noise sources and extended those benefits to additional spatial configurations.

All evaluations of the MOC strategy reported so far, however, were restricted to using time-interleaved but otherwise identical electrical pulse sequences across frequency channels, an approach termed continuous interleaved sampling (CIS) (Wilson et al., 1991). Some state-of-the-art CIs, however, deliver channel-specific electrical pulse sequences intended to preserve the temporal fine structure (TFS) cues in speech (*e.g.*, Zierhofer, 2001; Riss et al., 2008; Schatzer et al., 2010). One such strategy, featured in MED-EL clinical devices, is termed FS4 because it preserves the stimulus fine structure in the four most-apical frequency channels (Hochmair et al., 2006; Lorens et al., 2010; Schatzer et al., 2010; Riss et al., 2014). To assess the potential of MOC processing in an eventual implementation of the strategy in contemporary devices, one aim of the present study was to compare speech-in-noise recognition with combined MOC3 and FS4 processing relative to using FS4 processing alone. These two approaches are referred to hereafter as MOC3-FS4 and STD-FS4 strategies, respectively. The two strategies were implemented with identical automatic gain control (AGC). The AGC was such that the two processors in the pair applied equal broadband gain in the two ears, an approach sometimes referred to as 'linked' AGC (*e.g.*, Wiggins and Seeber 2013). Although this type of AGC is not standard in clinical CI devices (Potts et al., 2019; Archer-Boyd and Carlyon, 2019), it could be easily implemented in a binaural CI device and would theoretically preserve head-shadow interaural level differences (ILDs), which are useful for contralateral MOC processing to function properly (Lopez-Poveda et al., 2016b). Evaluations involved comparing SRTs for seven BiCI users with the two strategies for sentences in competition with a single source of steady-state speech-shaped noise (SSN) or an international female fluctuating masker (iFFM) (Holube et al., 2010), and for sentences presented at various sound levels.

On the other hand, although it has been shown that the MOC strategy can facilitate speech-in-noise intelligibility in some conditions (Lopez-Poveda et al., 2016b, 2017, 2020), it is not yet known to what extent this strategy affects listening effort (if at all). Listening effort is frequently defined as "the mental exertion required to attend to, and understand, an auditory message" (McGarrigle et al., 2014). CI users may need to exert less effort when listening with the STD-FS4 than with the MOC3-FS4 strategy because the STD-FS4 strategy is closer to the audio processing strategy in their clinical devices. On the other hand, however, CI users may need to exert less effort when listening with the MOC3-FS4 than with the STD-FS4 strategy because the MOC3-FS4 strategy can facilitate the recognition of speech in noise. A second aim of the present study was to compare listening effort with the MOC3-FS4 and STD-FS4 strategies. We hypothesized that listening with the MOC3-FS4 strategy requires the same or less effort as listening with the STD-FS4 strategy, indicating that the better intelligibility in noise with the MOC3-FS4 than with the STD-FS4 strategy is not the result of participants spending more effort with the MOC3-FS4 strategy. The concept of listening effort is complex and it is yet unclear how to best measure effort (for a review, see Pichora-Fuller et al. 2016). Here, we used two proxies for listening effort: the word recall score in a dual-task (word recognition and word recall) test (Gosselin and Gagne, 2011; Pichora-Fuller et al., 2016; Gagné et al., 2017) and the response time in a word recognition test (Gatehouse and Gordon, 1990; Houben et al., 2013; Gustafson et al., 2014; Meister et al., 2018).

2. Material and methods

2.1. Participants

Seven bilateral users of MED-EL CIs participated in the study (**Table 1**). All participants reported to perform very well with their implants. Participants were volunteers and not paid for their services. They all signed an informed consent to participate in the study. All participants were native speakers of Castilian Spanish. None of them had been previously tested in the laboratory with any of the sound processing strategies used in the study.

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

2.2. Processing strategies

Stimuli were processed through the STD-FS4 and MOC3-FS4 sound processing strategies. The two strategies were identical except for the back-end compression stage, as described below.

2.2.1. STD-FS4 strategy

In contrast to the reference STD strategy used in previous studies (Lopez-Poveda et al., 2016b, 2017, 2020), which used time-interleaved but otherwise identical electrical pulse sequences across frequency channels (Wilson et al., 1991), the STD-FS4 used here intended to preserve TFS cues in speech (*e.g.*, Zierhofer 2001, Riss et al. 2008, Schatzer et al. 2010). Furthermore, the STD-FS4 strategy was implemented with linked AGC, meaning that the AGC functions at the two ears applied identical gain, equal to the minimum gain across the ears (Wiggins and Seeber, 2013).

After the AGC, the STD-FS4 strategy involved two functionally independent sound processors, one per ear. Each processor included the following stages. (1) A bank of MED-EL's proprietary finite-impulse-response bandpass filters with a modified logarithmic distribution between 70 and 8500 Hz. (2) Envelope extraction via Hilbert transform. (3) A channel-specific gain to the input signal to the compression function. This gain replaced the high-pass pre-emphasis filter employed in previously tested implementations of the STD strategy (*e.g.*, Lopez-Poveda et al. 2016b, 2017, 2020). (4) Sampling of compressed envelopes with biphasic electrical pulses using the FS4 approach, *i.e.*, using channel-specific sampling sequences in the four most apical channels and time-interleaved fixed-rate stimulation sequences in the remaining channels [for more details about the FS4 strategy, see Riss et al. (2014)]. The number of filters in the filter banks 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. (Note that having an equal number of channels in the two ears was nec-

Table 1

Participants' data. 'Better ear' indicates the better hearing ear as reported by the participant. *c*: compression parameter value in the participants clinical CIs; Ch: cholesteatoma; *F*: female; Inf: infections; *L*: left; *M*: male; MCL: maximum comfortable loudness; Mg: meningitis; Nn: neurinoma; Os: otosclerosis; pps: pulses per second per electrode; *R*: right; Syn: syndromic; Un: unknown; Vol: volume. Note that the pulse rate was used only in the channels without FS4 processing. FS channels typically use a higher pulse sampling rate for a more accurate representation of the temporal fine structure.

ID	Sex	Age (years)	Etiology	Time of implant use (months)		Processor/Implant/Electrode array in the clinical devices		Processing strategy in the clinical devices		Electrodes Active/Used for testing		Pulse rate (pps)		Better ear	*c* value in the clinical devices		THR (%MCL)		Vol (%) STD MOC	
				L	R	L	R	L	R	L	R	L	R		L	R	L	R	L	R
021	F	40	Inf	29	187	RONDO CONCERTO FLEX28	RONDO CONCERTO FLEX28	FSP	FS4-p	1-12 1-12	1-12 1-12	1399	1210	L	500	500	10	10	85 95	75 85
022	M	49	Mg	185	200	OPUS2 C40+ Standard	RONDO SONATAti100 FLEXsoft	HDCIS	FS4-p	1-9 1-9	1-7,9-11 1-7,9-10	1500	1322	L	900	500	10	10	90 95	75 80
023	F	68	Os	211	184	OPUS2 C40+ Standard	OPUS2 C40+ Standard	FSP	FS4	1-4,6-10,12 1-4,6-10,12	1-11 1-10	1500	912	L	500	500	10	10	85 85	85 85
024	M	62	Nn/Ch	119	96	OPUS2 SONATAti100 Standard	OPUS2 SONATAti100 Standard	FS4-p	FS4-p	1-9 1-9	1-12 1-9	1268	1449	L	500	500	10	10	90 90	80 80
025	F	16	Un	169	180	OPUS2 PULSARci100 Standard	OPUS2 PULSARci100 Standard	FS4-p	FS4-p	1-9 1-8	1-8 1-8	1266	1293	R	500	500	10	10	90 95	90 95
026	M	19	Un	52	69	OPUS2 CONCERTO FLEXsoft	OPUS2 CONCERTO FLEXsoft	FS4-p	FS4-p	1-11 1-11	1-11 1-11	1258	1382	R	500	500	10	10	95 100	90 95
027	F	19	Syn	44	19	RONDO CONCERTO FLEXsoft	RONDO CONCERTO FLEXsoft	FS4-p	FS4-p	1-11 1-11	1-12 1-11	1302	1240	R	500	500	10	10	85 90	90 100

essary for the MOC strategy to exert its contralateral on-frequency inhibition, as described below).

The back-end compression function (or acoustic-to-electric map) in all processors was as follows (Boyd, 2006):

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

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. In the STD-FS4 strategy, the value of 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 (shown in **Table 1**).

Note that the STD-FS4 strategy was the most similar to those employed by the participants in their clinical devices (**Table 1**), except for the use of a linked AGC. We, however, used a research implementation of the FS4 strategy that could be slightly different from the implementation of FS4 strategy in the clinical audio processors.

2.2.2. *MOC3-FS4 strategy*

The MOC3-FS4 strategy was like the STD-FS4 strategy in all respects except that the value of the compression parameter (c in Eq. (1)) in every frequency channel of processing varied dynamically depending upon the time-weighted output level from the corresponding frequency channel in the contralateral processor (see Fig. 1 in Lopez-Poveda et al. 2016b). 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) (Fig. 2 in Lopez-Poveda et al. 2016b). The contralateral output level was calculated over an exponentially decaying time window with two time constants ($\tau_a = 2$ ms, $\tau_b = 300$ ms), and it was normalized to the channel bandwidth to achieve relatively greater inhibition in the apical than in the basal frequency channels. In other words, in the MOC3-FS4 strategy, the contralateral control of compression was identical as for the MOC3 strategy (without FS4 or AGC) described in Lopez-Poveda et al. (2019, 2020), where further details can be found.

2.3. *Fitting and loudness level balancing*

Before any 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 10% of MCL values (Boyd, 2006) (**Table 1**). Processor volumes were set independently for each ear and for each processing strategy (STD-FS4 and MOC3-FS4) to ensure that a sentence from a source at 0 ° azimuth and 0 ° elevation was perceived as comfortable, in the center of the head, and equally loud for the two strategies. Volumes were fitted separately for the STD-FS4 and MOC3-FS4 strategies (**Table 1**) in an attempt to compensate for the possible reduction in loudness associated with MOC3 processing (Lopez-Poveda et al., 2016b; Lopez-Poveda and Eustaquio-Martín, 2018). Once set, thresholds, MCL levels, and volumes remained constant for each participant across test conditions.

2.4. *Equipment and virtual acoustics*

The MATLAB software environment (R2015b, 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 (RIB2; Department of Ion Physics and Applied Physics at the University of Innsbruck, Innsbruck, Austria) and each patient's implanted receiver/stimulator(s).

Spatial configurations were achieved by convolving monophonic recordings with head-related impulse responses (HRIRs) obtained with the front microphone of a behind-the-ear MED-EL SONNET sound processor placed on an artificial head and torso (G.R.A.S. 45BB-2). Responses were recorded in a nearly-anechoic chamber (RT60 = 0.06 s, DRR = 13 dB). The levels of the stimuli were determined using the broadband root-mean-square amplitude and were set before HRIR convolution to preserve head-related cues.

2.5. *Speech reception thresholds*

2.5.1. *Procedure*

The procedure was identical to that of Lopez-Poveda et al. (2020). Intelligibility in noise was assessed by measuring the SNR at which participants 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 and varying the noise level adaptively using a one-down, one-up procedure (Levitt, 1971). For each SRT measurement, thirty sentences were presented, and participants were asked to repeat each sentence. A sentence was scored as correct when all of its words were recognized, and incorrect when one or more words were not recognized. The first ten sentences were always the same (taken from a practice list) irrespective of the test condition and for all participants but were presented in random order. This gave participants the opportunity to become familiar with the processing strategy tested during the corresponding SRT measurement. The initial SNR was 20 dB and it 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). Three SRTs were measured in this way for each test condition and the mean was regarded as the final SRT. If the standard deviation of the three estimates was greater than 6 dB, a fourth estimate was measured and included in the mean.

During the experiment, the presentation of each sentence was controlled by the experimenter. Participants were instructed to repeat what they heard, and the experimenter scored each sentence as correct or incorrect before presenting the next sentence. Feedback was not given to participants on the correctness of their responses.

2.5.2. *Test conditions*

Speech reception thresholds in noise were measured in bilateral listening (*i.e.*, listening with the two CIs) and in unilateral listening with the self-reported better ear alone (**Table 1**). In all cases, the speech and noise sources were at eye level (0 ° elevation). 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 different for different participants depending on the self-reported better ear of each participant. However, for convenience, results are reported as if the better ear was the right ear for all participants. The noise was always in front or toward the self-reported worse ear. In bilateral listening for speech at −38 dB FS, SRTs were measured for five spatial configurations of the target and masker sources (S_0N_{-90}, S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$, and $S_{90}N_{-90}$). In bilateral listening for speech level at −28 and −48 dB FS, SRTs were measured for three target-masker spatial configurations (S_0N_0, $S_{15}N_{-15}$, and $S_{90}N_{-90}$). In unilateral listening with speech at −38 dB FS, SRTs were measured for three target-masker spatial configurations (S_0N_0, $S_{15}N_{-15}$, and $S_{60}N_{-60}$). In the S_XN_Y notation, X and Y indicate the azimuthal angles (in degrees) of the speech (S) and noise (N) sources with 0 ° indicating a source directly in front and positive and negative

Table 2
Conditions and stimuli used to test each participant. HINT: hearing-in-noise test sentences. Matrix: matrix sentences. SSN: speech-shaped noise for HINT sentences. Matrix noise: speech-shaped noise for matrix sentences. iFFM: international female fluctuating masker. n.m.: not measured.

Masker type		SSN				iFFM
Speech level (dB FS)		−38	−28	−48	−38	−38
Listening mode		Bilateral	Bilateral	Bilateral	Unilateral	Bilateral
Participant						
021	*Speech*	HINT	HINT	HINT	n.m.	HINT
	Noise	SSN	SSN	SSN		iFFM
022	*Speech*	Matrix	Matrix	Matrix	Matrix	Matrix
	Noise	Matrix noise	Matrix noise	Matrix noise	Matrix noise	iFFM
023	*Speech*	Matrix	Matrix	Matrix	Matrix	Matrix
	Noise	Matrix noise	Matrix noise	Matrix noise	Matrix noise	iFFM
024	*Speech*	HINT	HINT	HINT	HINT	HINT
	Noise	SSN	SSN	SSN	SSN	iFFM
025	*Speech*	Matrix	Matrix	n.m.	n.m.	Matrix
	Noise	Matrix noise	Matrix noise			iFFM
026	*Speech*	Matrix	Matrix	Matrix	n.m.	Matrix
	Noise	Matrix noise	Matrix noise	Matrix noise		iFFM
027	*Speech*	HINT	HINT	HINT	HINT	HINT
	Noise	SSN	SSN	SSN	SSN	iFFM

values indicating sources to the right and the left of the midline, respectively.

2.5.3. Order of testing

Test conditions were administered in the same order for all participants, as follows:

Bilateral listening with speech level at −38 dB FS and with the SSN masker.
Bilateral listening with speech level at −38 dB FS and with the iFFM masker.
Bilateral listening with speech level at −28 dB FS and with the SSN masker.
Bilateral listening with speech level at −48 dB FS and with the SSN masker.
Unilateral listening with speech level at −38 dB FS and with the SSN masker.

For each of the five test conditions, measurements were organized in three blocks (one block per SRT estimate) and within each block, target-masker spatial configurations and processing strategies were administered in random order. In bilateral listening with speech at −38 dB FS (and with SSN or iFFM), each block involved measuring 10 SRTs (2 strategies × 5 spatial configurations). In bilateral listening with speech at −28 and −48 dB FS and in unilateral listening with speech at −38 dB FS, each block involved measuring 6 SRTs (2 strategies × 3 spatial configurations). Recall that three SRTs were obtained per condition. Therefore, a total of 114 SRTs were measured per participant, except for participants SA021 and SA026 for whom SRTs in unilateral listening were not measured and participant SA025, who was not tested in unilateral listening or in bilateral listening at −48 dB FS due to lack of time (see **Table 2**).

2.5.4. Stimuli

The full protocol involved measuring 114 SRTs, each of which required passing 20 test sentences (aside from the 10 initial practice sentences). Therefore, the protocol required having access to 2280 different test sentences to prevent potential confounding effects of sentence repetition from affecting the results. An attempt was made to test all participants using the female sentences in the Spanish version of the Oldenburg Sentence Test, also known as the ‘matrix’ test (Hochmuth et al., 2012). Matrix sentences have an identical syntactical structure (name, verb, numeral, noun, and adjective) and are formed by randomly selecting one out of the 10 different words that are available for each category (*i.e.* 50 words in total). In this way, up to 10^5 different test sentences can be created, that is many more sentences that were necessary. (The test includes a practice list with 10 sentences.) Additional advantages of matrix sentences are that they are phonetically balanced, are semantically unpredictable based on context, have low redundancy, and participants can be shown the 50 words in advance. Because of all this, the use of matrix sentences was preferred to minimize the potential confounding effects of participants learning the test sentences.

Some participants, however, could not recognize the matrix sentences, even in quiet and after several opportunities. Those participants were tested using the sentences for a male speaker of the Castilian Spanish version of the Hearing-in-Noise Test (HINT) (Nilsson et al., 1994; Huarte, 2008). Unlike matrix sentences, HINT sentences vary in length (they have a different number of words) and have context. The corpus of HINT sentences includes one practice list and 20 test lists with 10 sentences per list, *i.e.*, 200 different test sentences in total. Because this is many fewer sentences than were needed, HINT sentences had to be used multiple times for a given participant. To minimize the effects of sentence repetition from confounding the main sought effects of processing strategy or the potential interaction between strategy and spatial configuration, sentences were not repeated within a testing block. Recall that a testing block involved measuring one SRT estimate per processing strategy per spatial configuration in a given test condition, with strategies and spatial configurations administered in random order. For example, the first testing block was aimed at obtaining the first set of SRT estimates for condition #1 (Bilateral listening with speech level at −38 dB FS and with the SSN masker) for the two strategies (STD-FS4 and MOC-FS4) in five spatial configurations (S_0N_{-90}, S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$, and $S_{90}N_{-90}$). Therefore, this testing block involved measuring 10 SRTs, which required 200 test sentences, hence there was no need to repeat sentences within this testing block. The second testing block was aimed at obtaining a second set of SRT estimates for the same condition. Therefore, all HINT sentences had to be used again for the second testing block, and so on for subsequent blocks. For any testing block, however, sentences, strategies, and spatial configurations were administered in random order to minimize sentence repetition from affecting the main sought effects of strategy, spatial configuration, or their interaction. The effects of sentence repetition are further discussed later.

The sentence material used to test each participant is shown in **Table 2**. The use of different sentence materials for different participants was deemed reasonable because the aim was to compare performance across two processing strategies tested with the same speech material (within-subject comparison design), rather than to compare performance across participants. In other words, any effect of speech material was assumed to affect the two strategies equally for a given subject.

Speech reception thresholds were measured for sentences masked by SSN and an iFFM. The SSN was different for matrix and HINT sentences so that the SSN spectrum matched the average spectrum of the corresponding sentence material. A different SSN or iFFM token was used to mask each sentence. The masker started 500 ms before the sentence onset and ended 100 ms after the sentence offset and was gated with 50 ms cosine-squared onset and offset ramps. For the SSN masker, SRTs were measured for speech at −48, −38 and −28 dB FS. For the iFFM, the speech level was fixed at −38 dB FS. For reference, the speech level of −38 dB FS corresponds approximately to 65 dB SPL in MED-EL clinical CI audio processors.

2.6. Proxies for listening effort

We used two different methodologies to assess listening effort.

2.6.1. Dual-task paradigm: word recognition and recall

Participants were instructed to recognize and repeat each of 10 disyllabic words (primary task) and to remember the words for later recall (secondary task). The words were selected from the corpus of Cárdenas and Marrero (1994), a standard for clinical testing in Spain[1]. Words uttered by a male talker were presented in quiet or in competition with HINT SSN noise. The words were presented at a fixed level of −38 dB FS. All tests involved bilateral stimulation. Participants had to repeat each word after they heard it. A word was counted as correctly recognized when it was identical to the word presented. Feedback was not given to the participants on the correctness of their responses. As soon as the 10 words were presented, the participant was asked to recall as many words as he/she could remember, regardless of the order of presentation. Two scores were obtained: the number of correctly recognized words and the number of correctly recalled words. We assumed that the proportion of recalled words relative to the total number of presented words ($N = 10$) informs about the amount of effort spent by the listener in the recognition task (Rabbitt, 1966; Pichora-Fuller et al., 1995; Sarampalis et al., 2009). Three measurements were made per participant and per condition, and the mean was taken as the final score.

Word recognition and recall were measured at three SNRs: in quiet, at + 5 dB SNR, and at the individual SRTs for sentences in SSN (see **Table 3**). We chose to include a measure in quiet as a reference (practice) condition to give participants the opportunity to familiarize with the task. We chose + 5 dB SNR because it is a typical SNR in natural listening situations (Smeds et al., 2015). We measured effort at the individual SRT to investigate the effect of the processing strategy on listening effort in noise in conditions of equal intelligibility. Effort typically decreases with increasing intelligibility and for a fixed SNR, the MOC3-FS4 strategy could theoretically improve intelligibility in noise. Therefore, we expected MOC processing to decrease or not change listening effort for speech at a fixed SNR. However, we expected listening with MOC strategy to be as effortful as listening with the STD strategy in conditions of equal intelligibility.

Table 3
SRTs (in units of dB SNR) in bilateral listening for sentences at −38 dB FS in competition with SSN. Values are shown for each participant, strategy (STD-FS4 and MOC3-FS4), and spatial configuration ($S_{15}N_{-15}$ and $S_{60}N_{-60}$). Also shown are the group mean values. These values correspond to those shown in **Fig. 1**. s.d.: standard deviation.

		Strategy	
Condition	Participant	STD-FS4	MOC3-FS4
$S_{15}N_{-15}$	021	2.6	−4.0
	022	1.5	0.4
	023	1.8	−1.7
	024	6.9	4.9
	025	−3.5	−4.7
	026	−3.5	−3.3
	027	2.5	1.4
	Mean	1.2	−1.0
	s.d.	3.7	3.4
$S_{60}N_{-60}$	021	−0.1	−4.0
	022	1.3	−3.3
	023	−6.7	−5.3
	024	2.0	−0.3
	025	−9.6	−9.7
	026	−9.1	−6.6
	027	1.4	−3.0
	Mean	−2.9	−4.6
	s.d.	5.2	2.9

2.6.2. Verbal response time

The response time in the word recognition and recall task was used as another proxy for listening effort. For this purpose, participants wore a microphone while performing the word recognition and recall task, and their verbal responses were recorded for later scoring of response times using Adobe Audition v3.0. Note that participants were *not* instructed to give their response as quickly as possible. Response times were manually measured as the time elapsed from the offset of each word to the onset of the participant's response during the primary word-recognition task. Participants were instructed to always respond (*e.g.*, with 'next' or 'no') even if they did not recognize the word. Response times were measured for each of the 10 words that were presented in each test condition, regardless of whether the word was recognized or not. Sounds indicating hesitation or thinking were not regarded as a response in computing response times. Three measurements were made per participant and test condition, and the mean was taken as the final score. Longer response times were interpreted as reflecting greater effort in the word recognition task.

2.6.3. Test conditions

The two proxies for listening effort were measured in bilateral listening for two target-masker spatial configurations ($S_{15}N_{-15}$ and $S_{60}N_{-60}$), two processing strategies (STD-FS4 and MOC3-FS4), and three SNRs (quiet, + 5 dB, and the individual SRT). For each test condition, effort was assessed three times. This amounted to 36 effort estimates in total (2 spatial configurations × 2 strategies × SNRs × 3 estimates per condition).

The target was always presented at an azimuth ipsilateral to the listener̂s self-reported better ear, but the $S_{15}N_{-15}$ or $S_{60}N_{-60}$ nomenclature was chosen by convention. In the masked conditions, freshly generated SSN tokens were used to mask each word (*i.e.*, frozen noise was not used). The noise started 500 ms before the word onset and ended 100 ms after the word offset.

2.6.4. Order of testing

Effort was assessed first in quiet, followed by the + 5 dB SNR condition, and finally followed by the condition at the individual SRTs. For each of the three SNRs, conditions (spatial configurations

---

[1] The corpus is structured in lists of 25 words. Each list is phonetically balanced. Because all participants were adults, we used the lists for adults testing. The 10 words used for the present tests were selected randomly from each list of 25 words. A different list was selected at random and used for each measurement of effort.

and processing strategies) were administered in random order. Participants were given a brief break between test blocks.

2.7. Double-blind approach

All tests were 'double blind' such that neither the experimenter nor the participant knew which sound-processing strategy was being tested at any time.

2.8. Statistical analyses

Statistical analyses were conducted using IBM SPSS Statistics version 23.

Speech reception thresholds results. The results from unilateral and bilateral listening tests were analyzed separately. Kolmogorov-Smirnov tests (with Lilliefors corrections) were used to verify that the distributions of SRTs were normal (Gaussian). Two-way repeated-measures analyses of the variance (RMANOVA) were conducted to test for the effects of processing strategy, spatial configuration, and their interaction on group-mean SRTs. The Greenhouse-Geisser correction was applied when the sphericity assumption was violated. For tests involving multiple groups or variables, post hoc pairwise comparisons were conducted using Bonferroni corrections of the p value for multiple comparisons. We applied two-tailed tests for all analyses, and a result was regarded as statistically significant when $p \leq 0.05$.

Listening effort results. We obtained the proportion of recognized words and the proportion of recalled words. An arcsine transformation[2] was applied to the proportions of recognized and recalled words to make them suitable for further statistical analyses (Studebaker, 1985; Studebaker et al., 1995):

$$T[\mathrm{AU}] = \arcsin\left(\sqrt{\frac{s}{N+1}}\right) + \arcsin\left(\sqrt{\frac{s+1}{N+1}}\right), \qquad (2)$$

where s denotes the number of correct responses, N is the total number of trials (10 in this case), and T is the transformed proportion in arcsine units (AU).

Because three estimates of recognized and recalled words were obtained per condition, a transformed proportion (in AU units) was calculated for each of the three estimates and the mean was taken as the final transformed proportion.

Data were analyzed separately for the two spatial configurations tested ($S_{15}N_{-15}$ and $S_{60}N_{-60}$). If the distributions of recognized words, recalled words, and response times were normal (Gaussian), parametric RMANOVAs were used to test for the effects of processing strategy, SNR, and their interaction on the transformed proportions of recognized and recalled words. When the distributions were not normal, nonparametric Friedman tests were applied instead to test for the effect of test condition (given by processing strategy and SNR) and Wilcoxon signed-rank tests were applied post hoc for pairwise comparisons. Bonferroni correction for multiple comparisons was applied. An effect was regarded as statistically significant when the null hypotheses could be rejected with 95% confidence ($p \leq 0.05$). In the response time analysis, we report data from the full data set (*i.e.*, results are based on both recognized and not recognized words).

3. Results

3.1. Speech reception thresholds

3.1.1. SRTs in bilateral listening in steady noise

This section illustrates the benefits from MOC3-FS4 processing in bilateral listening for speech in SSN at different speech levels. The top row in **Fig. 1** (panels A–C) shows SRTs in bilateral listening with the STD-FS4 strategy for speech at −48, −38 and −28 dB FS, respectively, as indicated at the top of the panel, and for different spatial configurations (S_0N_{-90}, S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$ and $S_{90}N_{-90}$). In all panels, circles illustrate individual data and grey bars illustrate group mean scores ($N = 6$ at −48 dB FS, and $N = 7$ at −38 and −28 dB FS). Recall that each individual score is the mean of three estimates. The bottom row (panels D–F) illustrates SRT improvements (in dB) provided by MOC3-FS4 strategy relative to the STD-FS4 strategy. Positive values indicate that SRTs were lower (better) with the MOC3-FS4 than with the STD-FS4 strategy while negative values indicate that the MOC3-FS4 strategy was disadvantageous compared to the STD-FS4 strategy.

For speech at −48 dB FS (Fig. 1A,D), mean SRTs across participants were equal or better with the MOC3-FS4 strategy than with the STD-FS4 strategy for all spatial configurations. A two-way RMANOVA was conducted to test for the effect of processing strategy (STD-FS4 and MOC3-FS4), spatial configuration (S_0N_0, $S_{15}N_{-15}$ and $S_{90}N_{-90}$), and their interaction on the group mean SRTs. The RMANOVA revealed a significant effect of spatial configuration [$F(2,10) = 26.326$, $p < 0.001$]. However, the effect of processing strategy [$F(1,5) = 2.600$, $p = 0.168$] and the interaction between strategy and spatial configuration [$F(2,10) = 0.043$, $p = 0.958$] were not significant. The mean SRTs across participants and processing strategies were −1.1, −2.6 and −5.1 dB SNR for S_0N_0, $S_{15}N_{-15}$ and $S_{90}N_{-90}$, respectively. A pairwise post hoc analysis with Bonferroni correction for multiple comparisons revealed that SRTs were statistically better for spatial configurations where the speech and noise sources were spatially separated (S_0N_0 versus $S_{90}N_{-90}$, $p = 0.004$; and $S_{15}N_{-15}$ versus $S_{90}N_{-90}$ $p = 0.016$).

For speech at −38 dB FS (Fig. 1B,E), the MOC3-FS4 strategy tended to be advantageous on average over the STD-FS4 strategy for all spatial configurations except S_0N_0, where mean SRTs were approximately equal for the two strategies. The RMANOVA revealed significant main effects of processing strategy [$F(1,6) = 8.684$, $p = 0.026$] and spatial configuration [$F(4,24) = 28.583$, $p < 0.001$] on group mean SRTs. However, the interaction between strategy and spatial configuration was not statistically significant [$F(4,24) = 0.970$, $p = 0.442$]. The grand mean SRT was significantly better for the MOC3-FS4 than for the STD-FS4 strategy (−1.3 versus 0 dB SNR, $p = 0.026$). On the other hand, mean SRTs (across strategies and participants) tended to improve with increasing the spatial separation between the speech and noise sources (mean SRTs were 1.0, 2.1, 0.1, −3.8 and −2.6 dB SNR for S_0N_{-90}, S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$ and $S_{90}N_{-90}$, respectively). Pairwise post hoc comparisons (with Bonferroni corrections) also revealed significant differences in SRTs across some spatial configurations (S_0N_{-90} versus $S_{60}N_{-60}$, $p = 0.001$; S_0N_0 versus $S_{15}N_{-15}$, $p = 0.040$; S_0N_0 versus $S_{60}N_{-60}$, $p = 0.001$; S_0N_0 versus $S_{90}N_{-90}$, $p = 0.006$; $S_{15}N_{-15}$ versus $S_{60}N_{-60}$, $p = 0.015$; and $S_{15}N_{-15}$ versus $S_{90}N_{-90}$, $p = 0.006$).

For speech at −28 dB FS (Fig. 1C,F), mean SRTs tended to be equal or better with the MOC3-FS4 strategy than with the STD-FS4 strategy for all spatial configurations. A RMANOVA revealed a significant main effect of spatial configuration [$F(2,12) = 39.346$, $p < 0.001$] on mean SRTs. The effect of processing strategy [$F(1,6) = 5.460$, $p = 0.058$] and the interaction between strategy and spatial configuration [$F(2,12) = 0.098$, $p = 0.907$], however, were not statistically significant. Pairwise post hoc comparisons

---

[2] In the speech perception literature, it is common to apply the 'rationalized' arcsine transform (Studebaker, 1985) and express the transformed proportions in rationalized arcsine units (RAUs) rather than the arcsine units (AUs). We, however, disregarded applying a rationalized arcsine transform because the rationalization is not accurate for proportion values less than 20% and higher than 80%, and we often found proportions larger than 80% in the present data set.

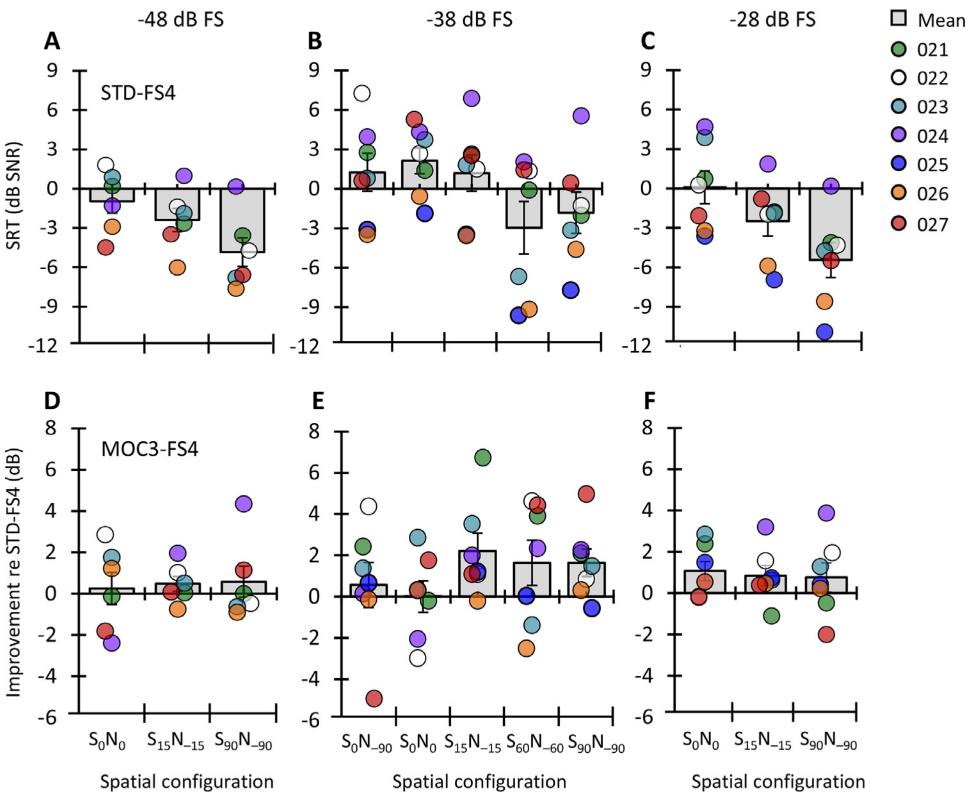

Fig. 1. Top (A–C). SRTs for sentences in competition with SSN in bilateral listening with the STD-FS4 strategy for speech at −48, −38 and −28 dB FS, respectively. **Bottom (D–F)**. SRT improvement provided by the MOC3-FS4 strategy relative to the STD-FS4 strategy. Circles illustrate individual data (as the average of three independent measures), grey bars illustrate group mean scores ($N = 6$ at −48 dB FS, and $N = 7$ at −38 and −28 dB FS), and error bars illustrate one standard error of the mean. In each panel, the abscissa is the spatial configuration of the speech and noise sources. See the main text for details.

(using Bonferroni correction) revealed that SRTs were significantly different for every pair of spatial configurations ($p < 0.05$), that is, SRTs tended to improve (become lower) with increasing the spatial separation between speech and noise sources (mean SRTs across participants and processors were −0.5, −2.9 and −5.8 dB SNR for S_0N_0, $S_{15}N_{-15}$ and $S_{90}N_{-90}$, respectively).

3.1.2. SRTs in unilateral listening in steady noise

In unilateral listening, SRTs were measured for four bilateral CI users, for speech at −38 dB FS in competition with SSN, and for spatial configurations of S_0N_0, $S_{15}N_{-15}$, and $S_{60}N_{-60}$. Results are shown in Fig. 2. A RMANOVA revealed a significant main effect of strategy [$F(1,3) = 19.717$, $p = 0.021$] and spatial configuration [$F(2,6) = 465.027$, $p < 0.001$]. The interaction between strategy and spatial configuration was not statistically significant [$F(2,6) = 0.400$, $p = 0.687$]. The mean SRT (across spatial configurations and participants) was 1.7 dB better for the MOC3-FS4 than for the STD-FS4 strategy (−5.4 versus −3.7 dB SNR, $p = 0.021$). Post hoc pairwise comparisons (with Bonferroni correction) revealed that SRTs were significantly different for every pair of spatial configurations ($p < 0.05$), namely, SRTs tended to improve with increasing the spatial separation between speech and noise sources (the mean SRTs across participants and strategies were −0.8, −3.7 and −9.1 dB SNR for S_0N_0, $S_{15}N_{-15}$ and $S_{60}N_{-60}$, respectively).

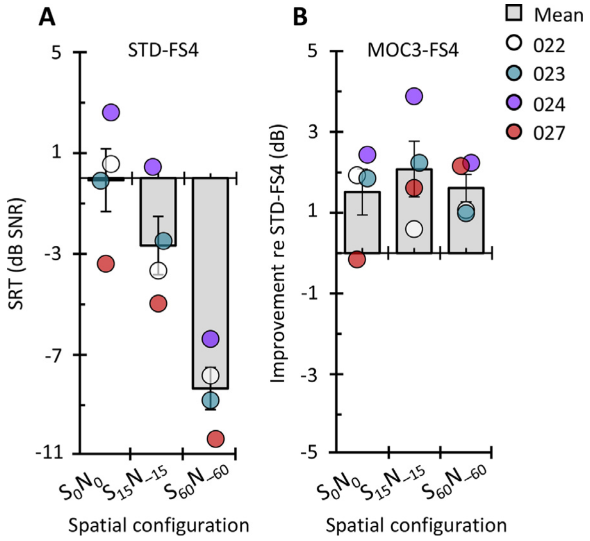

Fig. 2. A. SRTs with the STD-FS4 strategy in unilateral listening for speech at −38 dB FS in competition with SSN, and for three different spatial configurations of the speech and noise sources, as indicated at the bottom. **B.** SRT improvement with the MOC3-FS4 strategy relative to the STD-FS4 strategy. Error bars illustrate one standard error of the mean ($N = 4$).

3.1.3. SRTs in bilateral listening in fluctuating noise

For speech in competition with the iFFM, SRTs were measured in bilateral listening for speech at −38 dB FS. Results are shown in Fig. 3. A two-way RMANOVA revealed a significant effect of spatial configuration [$F(4,24) = 124.076$, $p < 0.001$] on SRTs, but no significant effect of processing strategy [$F(1,6) = 0.013$, $p = 0.913$] or interaction between strategy and spatial configuration [$F(4,24) = 1.176$, $p = 0.346$]. SRTs tended to be better with increasing the spatial separation of the speech and noise sources (mean SRTs across participants and strategies were 2.3, 6.6, 2.2, −2.4 and −0.3 dB SNR for S_0N_{-90}, S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$ and $S_{90}N_{-90}$, respectively). A pairwise post hoc analysis with Bonferroni correction for multiple comparisons revealed significant differences in mean SRTs (pooled across participants and strategies) between some spatial configurations (S_0N_0 versus S_0N_{-90},

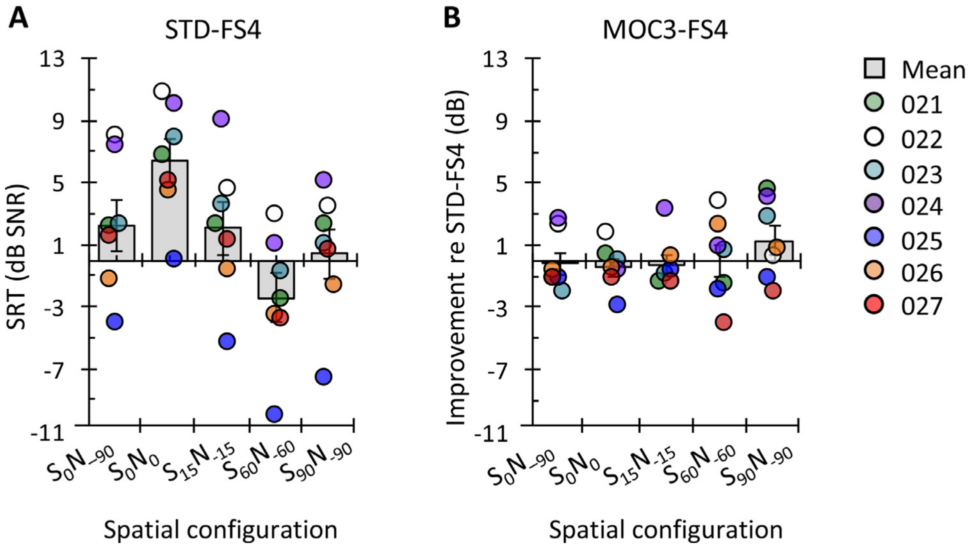

Fig. 3. **A.** SRTs with the STD-FS4 strategy in bilateral listening for speech at −38 dB FS in competition with an iFFM. **B.** SRT improvement with the MOC3-FS4 strategy relative to the STD-FS4 strategy. Error bars illustrate one standard error of the mean ($N = 7$).

$p < 0.001$; S_0N_0 versus $S_{60}N_{-60}$, $p < 0.001$; S_0N_0 versus $S_{90}N_{-90}$ $p = 0.018$; S_0N_{-90} versus $S_{15}N_{-15}$, $p = 0.001$; S_0N_{-90} versus $S_{60}N_{-60}$, $p < 0.001$; S_0N_{-90} versus $S_{90}N_{-90}$, $p < 0.001$; $S_{15}N_{-15}$ versus $S_{60}N_{-60}$, $p < 0.001$; $S_{15}N_{-15}$ versus $S_{90}N_{-90}$, $p = 0.045$; $S_{60}N_{-60}$ versus $S_{90}N_{-90}$, $p = 0.006$).

3.2. Proxies for listening effort

In this section, we first compare word recognition and word recall scores for the MOC3-FS4 and the STD-FS4 strategies. Then, we compare verbal response times for the two strategies. Lastly, we report a correlation analysis between word recall scores and response times.

3.2.1. Word recognition and recall

Fig. 4A illustrates the transformed proportion of *recognized* words (in AU units) in bilateral listening for the two processing strategies tested. Each bar is the mean score across seven bilateral CI users. Circles illustrate individual data; recall that the score for each participant and test condition was the mean of three estimates. Each column shows data for different spatial configuration ($S_{15}N_{-15}$ and $S_{60}N_{-60}$), as indicated at the top of the column.

For the $S_{15}N_{-15}$ spatial configuration, a two-way RMANOVA revealed no significant effects of processing strategy [$F(1,6) = 0.109$, $p = 0.753$], SNR [$F(2,12) = 3.591$, $p = 0.060$], or interaction between strategy and SNR [$F(2,12) = 0.224$, $p = 0.803$] on the transformed proportion of recognized words.

For the $S_{60}N_{-60}$ spatial configuration, a two-way RMANOVA showed that the effects of processing strategy [$F(1,6) = 0.006$, $p = 0.940$] and the interaction between strategy and SNR [$F(2,12) = 1.379$, $p = 0.289$] were not statistically significant. However, the effect of SNR was statistically significant [$F(2,12) = 5.367$, $p = 0.022$]. Post hoc pairwise comparisons with Bonferroni corrections revealed a significantly higher proportion of recognized words at + 5 dB SNR than at the individual SRT in noise ($p = 0.013$). However, we found no statistically significant differences between the proportion of recognized words for the other SNRs (quiet versus + 5 dB SNR, $p = 0.408$; quiet versus SRT, $p = 0.786$).

Fig. 4B shows the transformed proportion of *recalled* words (in AU units) in bilateral listening for the STD-FS4 and MOC3-FS4 strategies. For the $S_{15}N_{-15}$ spatial configuration, a two-way RMANOVA revealed that the effect of processing strategy [$F(1,6) = 0.845$, $p = 0.393$], the effect of SNR [$F(2,12) = 0.767$, $p = 0.486$], and the interaction between strategy and SNR [$F(2,12) = 0.809$, $p = 0.468$] were not significant. For the $S_{60}N_{-60}$ spatial configuration, a two-way RMANOVA showed that the effect of processing strategy [$F(1,6) = 2.711$, $p = 0.151$], the effect of SNR [$F(2,12) = 3.166$, $p = 0.079$], and the interaction between strategy and SNR [$F(2,12) = 0.259$, $p = 0.776$] were not statistically significant. Overall, we found no significant differences in word recall scores across sound-processing strategies or SNRs. In addition, there was no interaction between strategy and SNR for either of the two spatial configurations tested.

3.2.2. Verbal response time

Fig. 5 displays mean verbal response times for the two processing strategies (STD-FS4 and MOC3-FS4), for the three SNRs (quiet, + 5 dB SNR and individual SRT in noise), and for the two spatial configurations ($S_{15}N_{-15}$ and $S_{60}N_{-60}$). Each data point is the mean response time across seven BiCI users. Note that the score for each participant and test condition was the mean of three estimates.

Response times tended to increase with decreasing the SNR, i.e., response times tended to be shorter in quiet or at +5 dB SNR than for words presented at the individual SRT in noise. This is consistent with expectations because the individual SRTs were generally negative across conditions (Table 3) and hence word recognition tended to be harder (though not always significantly) at the individual SRTs in noise (Fig. 4A). In addition, response times were similar for the two processing strategies. Friedman tests revealed that response times were not statistically significantly different across the six test conditions (2 strategies × 3 SNRs), neither for the $S_{15}N_{-15}$ spatial configuration [$\chi^2(5) = 3.816$, $p = 0.576$] nor for the $S_{60}N_{-60}$ spatial configuration [$\chi^2(5) = 5.367$, $p = 0.373$]. In other words, we found no significant differences in verbal response times across sound-processing strategies or SNR for either spatial configuration.

3.2.3. Correlation between the two measures of listening effort

We conducted a correlation analysis to investigate if the two proxies for listening effort (word recall and verbal response time) reflected the same dimension. Fig. 6 shows a plot of the number

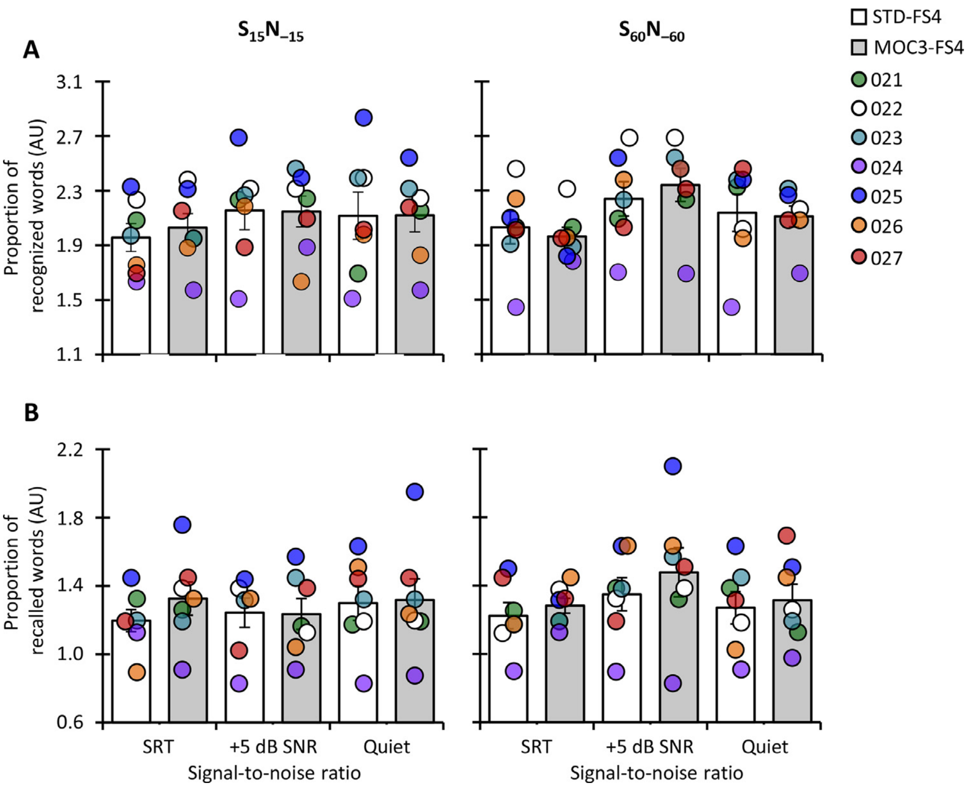

Fig. 4. Mean proportion of recognized (**A**) and recalled words (**B**) for the STD-FS4 (white bars) and MOC3-FS4 (grey bars) strategies, for the $S_{15}N_{-15}$ (**left**) and $S_{60}N_{-60}$ (**right**) spatial configurations and for the three SNRs. Circles illustrate individual mean scores for seven participants. Each circle is the mean of three measures. Error bars indicate one standard error of the mean ($N = 7$).

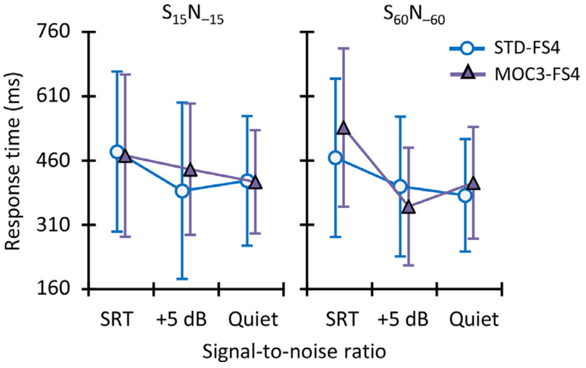

Fig. 5. Mean verbal response time in the word recognition task averaged across subjects ($N = 7$). The left and right panels illustrate response times for the $S_{15}N_{-15}$ and $S_{60}N_{-60}$ spatial configurations, respectively. Each panel illustrates response times for three different SNRs (abscissae) and two different processing strategies, as indicated by the inset. In the two panels, error bars illustrate one standard error of the mean ($N = 7$).

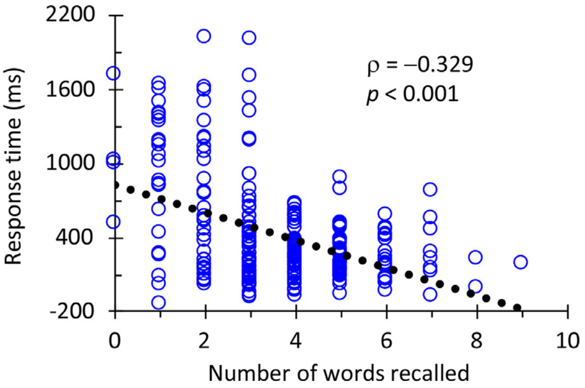

Fig. 6. Correlation between the verbal response time and the number of recalled words. Data are pooled across seven participants, two processing strategies (STD-FS4 and MOC3-FS4), two spatial configurations ($S_{15}N_{-15}$ and $S_{60}N_{-60}$), three different SNRs (quiet, +5 dB SNR and individual SRT in noise), and three estimates per test condition for each participant (252 points in total). The dotted line is a linear regression fit to the data. Also shown is Spearman's rank correlation coefficient (ρ) and the level of significance (p).

(rather than the transformed proportion) of recalled words against the verbal response time for each measurement. Note that the total number of data points in the figure (252) equals the number of participants ($N = 7$) times 12 test conditions (2 strategies × 3 SNRs × 2 spatial configurations) times three estimates per condition per participant. The figure reveals a trend for word recall to decrease with increasing response time. Spearman's rank correlation coefficient between the two variables was negative and statistically significant ($\rho = -0.329$, $p < 0.001$). This suggests that the two measures partially reflected the same dimension. In other words, to the extent that these metrics can be considered indicators of effort, a greater number of recalled words and shorter response times possibly reflected less listening effort.

4. Discussion

The aim of the present study was to experimentally compare intelligibility and listening effort with and without MOC processing in combination with linked AGC and FS4 processing.

4.1. The benefit from MOC processing at different speech levels

Although SRTs were equal or better with the MOC3-FS4 than with STD-FS4 strategy at the three speech levels tested, the benefit from MOC3-FS4 processing was overall greater for speech at −38 dB FS than at −48 or −28 dB FS (compare the bottom panels in Fig. 1). This is possibly because the MOC3 parameters were carried forward from the study of Lopez-Poveda et al. (2020), where parameters were optimized for a strategy without AGC or FS4 and for a speech level of −20 dB FS, which roughly corresponds to −38 dB FS when AGC is used. Both these levels (−20 dB FS without AGC and −38 dB FS with AGC) correspond approximately to 65 dB SPL in MED-EL clinical devices, which is a typical conversational level. Therefore, the present data suggest that MOC3 processing, as implemented here, provides a benefit for speech levels of about 65 dB SPL but could also provide some benefit, albeit smaller, for speech levels of 55 and 75 dB SPL.

4.2. The benefit from MOC processing for different masker types

Compared to the STD-FS4 strategy, the MOC3-FS4 strategy improved SRTs for sentences in SSN but did not improve or degrade SRTs for sentences in iFFM when tests were conducted at the same speech level (−38 dB FS) (compare Fig. 1E with Fig. 3B). At first sight, this finding might appear inconsistent with reported evidence that MOC processing improves speech recognition for speech in competition with a single-talker masker (Lopez-Poveda et al., 2017). MOC processing, however, was implemented differently in the two studies. The contralateral control of compression was slow in the present MOC3-FS4 strategy ($\tau_a = 2$ ms, $\tau_b = 300$ ms) but fast ($\tau_a = \tau_b = 2$ ms) in the study of Lopez-Poveda et al. (2017). While a slower contralateral control of compression is closer to the natural MOC reflex (e.g., Backus and Guinan 2006) and objectively improves the SNR in steady-state maskers (Lopez-Poveda and Eustaquio-Martín, 2018), a faster control of contralateral compression enhances the SNR in the ear contralateral to the masker (ipsilateral to the speech) when the masker is fluctuating (see Fig. 1 in Lopez-Poveda et al. 2020). Altogether, this suggests that the benefit from the MOC3-FS4 strategy for speech recognition in a fluctuating masker would be probably larger with a faster contralateral control of compression than was tested here.

4.3. Assessment of listening effort

We measured word recall scores and verbal response times on the assumption that they are reasonable proxies for listening effort. This assumption is supported by some trends in the present data. First, word recall scores tended to be better at + 5 dB SNR than at the individual SRT in noise, regardless of spatial configuration. This is probably because individual SRTs in noise were generally (very) negative for most participants (Table 3), and word recognition is harder at lower SNRs. [Note that, though not statistically significantly, word recognition and recall scores tended to be better at + 5 dB SNR than in quiet almost certainly because the quiet condition was administered first and participants were more experienced in the task at the time that it was administered at + 5 dB SNR]. Second, response times tended to increase with decreasing SNR (Fig. 5), which seems reasonable because speech recognition gets harder with increasing noise levels. Both these trends (poorer recall scores and longer response times with decreasing SNR) are consistent with the idea that increased task difficulty demands more cognitive load for the primary task (word recognition), leaving less cognitive resources for the secondary task (word recall) and increasing processing and response times. A third piece of evidence supporting our assumption is that the two metrics correlate with each other (Fig. 6), and the scores for either metric tend to get worse with increasing task difficulty (e.g., with increasing levels of background noise). This suggests that the two metrics partly reflect the same dimension, presumably effort. Admittedly, however, the effect of SNR did not reach statistical significance on any of the two metrics. Although this might be due to the small sample size ($N = 7$), it nevertheless undermines the validity of our assumption that word recall scores and response times are reasonable proxies for listening effort.

We hypothesized that recognizing speech in noise would require the same or less effort with the MOC3-FS4 than with the STD-FS4 strategy. We found no significant differences in the proportion of recalled words or in response times across the two strategies. If the proportion of recalled words and response time scores were a measure of effort, our results would indicate that BiCI users experienced similar amounts of effort when listening with the MOC3-FS4 and the STD-FS4 strategies. This would be a positive finding, as participants were almost certainly more accustomed to listening to speech processed through the STD-FS4 than the MOC3-FS4 strategy because the STD-FS4 strategy was closer to the audio coding strategy implemented in their clinical devices (Table 1). As explained in the preceding paragraph, however, the lack of a statistically significant effect of SNR on word recall scores and response times may reflect that these metrics and/or the experimental design were not sensitive enough to extract information related with listening effort.

4.4. Limitations

Test conditions were administered in the same order for all participants. Fig. 7 illustrates SRTs as a function of trial number for six participants. Clearly, SRTs improved gradually (became lower) with increasing trial number. Improvements occurred for participants tested with both HINT and matrix sentences, even though matrix sentences are specifically designed to minimize learning effects by omitting semantic information and by informing participants of all words in advance (Hochmuth et al., 2012; see Methods). This suggests that the SRT improvements illustrated in Fig. 7 were not (only) due to participants learning the sentences or their words and, instead, were due (at least partly) to participants learning how to perform the task (Schlueter et al., 2016). Insofar as speech recognition can improve with practice (Schlueter et al., 2016), administering conditions in the same ordered rendered comparisons of SRTs across conditions (e.g., speech level or listening modality) unreliable. For example, practice almost certainly explains why SRTs in bilateral listening and in SSN were higher (worse) for the first test condition (dark blue symbols in Fig. 7) than for the fourth test condition (purple symbols) even though the speech level was higher, thus more audible, in the first than in the fourth condition (−38 versus −48 dB FS, respectively). The numbers to the right of each panel show the SRT difference between those two conditions. In other words, SRTs were better at the lower speech level presumably because this was the fourth test condition and was administered towards the end of the testing week.

Three out of the seven participants were tested with HINT rather than matrix sentences (Table 2). Due to the limited number of HINT sentence lists, HINT lists had to be used multiple times to complete the protocol. For this reason, it is likely that participants tested with the HINT sentences learnt some of those sentences, or their words, during testing. As a result, the corresponding SRTs are probably better than they would have been if the speech material had not been used repeatedly. Re-using the sentences, however, unlikely contributed to the reported differences in SRTs across strategies (or spatial configurations or the interaction between strategy and spatial configuration) because anyone testing block involved testing processing strategies (and spatial configurations) in random order, before moving on to the next testing block. In other words, each testing block included SRT measurements that combined the two strategies and spatial configurations in random order (each group of symbols in Fig. 7 illustrates SRTs for three testing blocks, one per SRT estimate per condition). Therefore, the

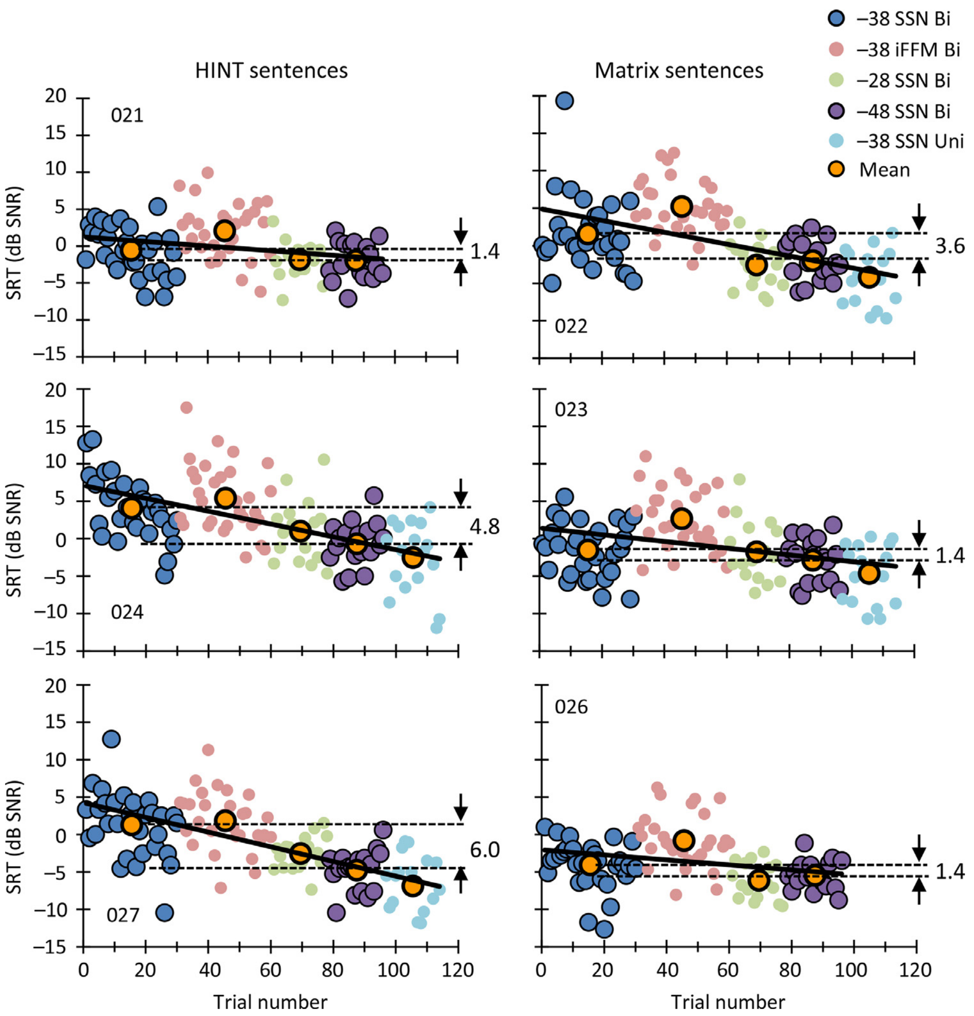

Fig. 7. SRTs as a function of trial number for six participants (021, 022, 023, 024, 026 and 027). Values are pooled for all tested conditions, including sound-processing strategies, target-masker spatial configurations, masker type, and speech level. Different colors illustrate groups of SRTs for a given test condition, defined by the speech level, masker type (SSN or iFFM) and listening modality: bilateral listening (Bi) or unilateral listening with the better ear (Uni), as indicated in the inset. The thick continuous lines are linear regression fits across all data points and orange symbols depict mean scores for each condition. Participants 022, 023 and 026 were tested using matrix sentences, while participants 021, 024 and 027 were tested using HINT sentences, as indicated at the top of each column. The numbers on the right-hand side of each panel show the difference between the mean SRTs for the first and the fourth test conditions (bilateral listening in SSN with speech level at −38 versus −48 dB FS, respectively). See the main text for details (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

learning of HINT speech material would have affected all strategies and spatial configurations similarly. The learning of the HINT speech material, however, likely contributed to the SRT improvements over time described in the preceding paragraph, which rendered comparisons of SRTs across conditions unreliable.

At the time when the tests were conducted, all participants had a long daily experience with an audio-coding strategy similar to the STD-FS4 (Table 1). By contrast, their experience with the MOC3-FS4 strategy was limited to a few days during the test sessions. For this reason, the found benefits of MOC processing for speech-in-noise recognition albeit small are encouraging. For CI users, speech recognition can improve significantly over time and with training (e.g., Dorman and Spahr 2006). For this reason, it is conceivable that the benefits from the MOC3-FS4 strategy could become larger with more practice and/or a sustained use of this strategy.

Given the lack of a consensus on the best measure of listening effort, we used two different methodologies (dual-task paradigm and verbal response time) to assess the effort experienced by CI users with the MOC strategy in a speech recognition task. The principal limitation of using the response time as a measure of effort is that it is not a ‘pure measure’ of effort, i.e., multiple aspects other than effort can influence the speed of processing, and hence response time, including age (Pichora-Fuller et al., 2016). A second limitation is that the response time might not always be sensitive to listening effort. For instance, a greater difficulty of the task could result in increased effort to maintain the same level of performance without observing differences in response time. Alternatively, it is possible that increased effort to maintain task performance may result in shorter response times (Bess and Hornsby, 2014). On the other hand, the dual-task paradigm, as a method for assessing listening effort, has ecological validity because it mimics some of the multitask difficulties that people face in real-world communication (Johnson et al., 2015). However, there is some uncertainty about whether measures of the behavioral consequences of listening in difficult environments (e.g. word recall scores) are a direct measure of mental effort (McGarrigle et al., 2014). The dual-task paradigm can be affected by individual differences in aspects such as task engagement and motivation (Alhanbali et al., 2019). In addition, behavioral dual-task paradigms are imprecise (Ohlenforst et al., 2017). The assumption that people use all their cognitive capacity to perform the primary and secondary task is not entirely accurate, since it is not possible to verify whether participants use all their cognitive capacity or not. Further, it is not possible to know with certainty if the participant always prioritizes the performance of the primary task (Alhanbali et al., 2019).

The test paradigm used here for assessing listening effort did not reveal significant effects of processing strategy or SNRs. The lack of effect of SNR can be interpreted as indicating that the chosen metrics were not sensitive enough to extract the information we were looking for. Future studies should use more sensitive metrics; perhaps, speech materials with more semantic context, in which cognitive abilities are more relevant, or objective metrics, such as, for example, pupillometry.

This study demonstrates that MOC3 processing combined with FS4 processing can produce equal or better SRTs in noise than a STD-FS4 strategy (all tested at the same speech level), without affecting word recall scores or verbal response times. This study, however, is only a first attempt to evaluate the potential benefits of MOC processing when implemented together with state-of-the-art sound-coding strategies. Further research is necessary to investigate the benefits of combining MOC processing with linked AGC, with pre-processing beamformers, and/or with other sound-coding strategies. It would also be important to test the MOC strategy in more realistic listening environments, including settings with reverberation and diffuse noise. In addition, because everyday hearing is dynamic (i.e., sound sources are mobile) and MOC processing is also dynamic, it would be appropriate to evaluate hearing performance with the MOC strategy using moving sound sources in real-world listening scenarios.

5. Conclusion

This study was aimed at comparing the possible benefits for speech-in-noise recognition of combining MOC3 with broadband linked AGC and FS4 processing relative to using FS4 processing with linked AGC alone. SRTs were compared for sentences presented in competition with steady (SSN) and fluctuating (iFFM) maskers and for various speech levels (−28, −38, and −48 dB FS) and spatial configurations. A second aim was to compare the processing strategies on word recall scores and verbal response times in a word-in-noise recognition task, two metrics often used as proxies for listening effort. The main conclusions are:

1 For BiCI users tested in bilateral listening, mean SRTs tended to be equal or better with the MOC3-FS4 than with the STD-FS4 strategy for all speech levels (−28, −38 and −48 dB FS) and maskers (SSN and iFFM). In steady noise, the mean SRT improvement across the range of speech levels and spatial configurations was 0.8 dB SNR. For speech at −38 dB FS, the mean SRT improvement across spatial configurations was 1.2 dB and the largest improvement was 2.2 dB SNR in the in $S_{15}N_{-15}$ spatial configuration. In fluctuating noise, SRTs were equal for the two strategies.
2 For bilateral CI users tested in unilateral listening, the mean SRT was statistically better with the MOC3-FS4 than with the STD-FS4 strategy. The mean SRT improvement across the tested speech-noise spatial configurations was about 1.7 dB.
3 Word recall scores and verbal response times were not statistically significantly different for the STD-FS4 and MOC3-FS4 strategies or across the tested SNRs (quiet, +5 dB SNR, and the individual SRT). It remains uncertain if this is because listening with the two strategies requires comparable effort and/or because word recall scores and verbal response times (as measured here) were insensitive to effort.

CRediT authorship contribution statement

Milagros J. Fumero: Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Almudena Eustaquio-Martín:** Methodology, Software, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **José M. Gorospe:** Resources, Writing – review & editing. **Rubén Polo López:** Resources, Writing – review & editing. **M. Auxiliadora Gutiérrez Revilla:** Resources, Writing – review & editing. **Luis Lassaletta:** Resources, Writing – review & editing. **Reinhold Schatzer:** Methodology, Writing – review & editing, Supervision. **Peter Nopp:** Supervision, Writing – review & editing, Project administration. **Joshua S. Stohl:** Methodology, Software, Resources, Writing – review & editing. **Enrique A. Lopez-Poveda:** Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgments

Work supported by MED-EL GmbH (Innsbruck, Austria), the Spanish Ministry of Science and Innovation (appropriate to note more clearly that all grants are to author EALP, Grant No. PID2019-108985GB-I00), and the European Regional Development Fund.

References

Alhanbali, S., Dawes, P., Millman, R.E., Munro, KJ., 2019. Measures of listening effort are multidimensional. Ear Hear. 40, 1084–1097.
Archer-Boyd, A.W., Carlyon, R.P., 2019. Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. J. Acoust. Soc. Am. 145, 1389–1400.
Backus, B.C., Guinan, J.J., 2006. Time-course of the human medial olivocochlear reflex. J. Acoust. Soc. Am. 119, 2889–2904.
Bess, F.H., Hornsby, BW., 2014. Commentary: listening can be exhausting-fatigue in children and adults with hearing loss. Ear Hear. 35, 592.
Boyd, PJ., 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.
Cárdenas, M.R., Marrero, V., 1994. Cuaderno de logoaudiometría. Universidad Nacional de Educación a Distancia (UNED), Madrid.
Dorman, M.F., Spahr, AJ., 2006. Speech perception by adults with multichannel cochlear implants. In: Waltzman, SB, Roland, JT (Eds.), Cochlear Implants. Thieme, New York, pp. 193–204.
Gagné, J.P., Besser, J., Lemke, U., 2017. Behavioral assessment of listening effort using a dual-task paradigm: a review. Trends Hear. 21, 1–25.
Gatehouse, S., Gordon, J., 1990. Response times to speech stimuli as measures of benefit from amplification. Br. J. Audiol. 24, 63–68.
Gosselin, P.A., Gagné, JP., 2011. Older adults expend more listening effort than young adults recognizing speech in noise. J. Speech, Lang. Hear. Res. 54, 944–958.
Gustafson, S., McCreery, R., Hoover, B., Kopun, J.G., Stelmachowicz, P., 2014. Listening effort and perceived clarity for normal hearing children with the use of digital noise reduction. Ear Hear. 35, 183–194.
Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Schösser, H., Garnham, C., Anderson, I., 2006. MED-EL cochlear implants: state of the art and a glimpse into the future. Trends Amplif. 10, 201–219.
Hochmuth, S., Brand, T., Zokoll, M.A., Castro, F.Z., Wardenaga, N., Kollmeir, B., 2012. A Spanish matrix sentence test for assessing speech reception thresholds in noise. Int. J. Audiol. 51, 536–544.
Holube, I., Fredelake, S., Vlaming, M., Kollmeier, B., 2010. Development and analysis of an international speech test signal (ISTS). Int. J. Audiol. 49, 891–903.
Houben, R., Van Doorn-Bierman, M., Dreschler, WA., 2013. Using response time to speech as a measure for listening effort. Int. J. Audiol. 52, 753–761.
Huarte, A., 2008. The Castilian Spanish hearing in noise test. Int. J. Audiol. 47, 369–370.
Hughes, K.C., Galvin, KL., 2013. Measuring listening effort expended by adolescents and young adults with a unilateral or bilateral cochlear implants or normal hearing. Cochlear Implant. Int. 14, 121–129.
Hughes, S.E., Hutchings, H.A., Rapport, F.L., McMahon, C.M., Boisvert, I., 2018. Social connectedness and perceived listening effort in adult cochlear implant users: a grounded theory to establish content validity for a new patient-reported outcome measure. Ear Hear. 39, 922–993.
Johnson, J., Xu, J., Cox, R., Pendergraft, P., 2015. A comparison of two methods for measuring listening effort as part of an audiologic test battery. Am. J. Audiol. 24, 419–431.
Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49, 467–477.

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. Acous. Soc. Am. 125, 372–383.

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

Lopez-Poveda, E.A., Eustaquio-Martín, A., Fumero, M.J., Stohl, J.S., Schatzer, R., Nopp, P., Wolford, R.D., Gorospe, J.M., Polo, R., Gutiérrez Revilla, M.A., Wilson, B.S., 2019. Lateralization of virtual sound sources with a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. Hear. Res. 379, 103–116.

Lopez-Poveda, E.A., Eustaquio-Martin, A., Stohl, J.S., Wolford, R.D., Schatzer, R., Wilson, BS., 2016b. A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. Ear Hear. 37, 138–148.

Lopez-Poveda, E.A., Eustaquio-Martín, A., Stohl, J.S., Wolford, R.D., Schatzer, R., Wilson, BS., 2016a. 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., Santa Cruz Ruiz, S., 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.

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, 2217–2231.

Lorens, A., Zgoda, M., Obrycka, A., Skarzynski, H., 2010. Fine structure processing improves speech perception as well as objective and subjective benefits in pediatric MED-EL COMBI 40+ users. Int. J. Pediatr. Otorhinolaryngol. 74 (12), 1372–1378.

McGarrigle, R., Munro, K.J., Dawes, P., Stewart, A.J., Moore, D.R., Barry, J.G., Amitay, S., 2014. Listening effort and fatigue: what exactly are we measuring? A british society of audiology cognition in hearing special interest group "white paper. Int. J. Audiol. 53, 433–440.

Meister, H., Rählmann, S., Lemke, U., Besser, J., 2018. Verbal response times as a potential indicator of cognitive load during conventional speech audiometry with Matrix sentences. Trends Hear. 22, 1–11.

Nilsson, M., Soli, S.D., Sullivan, JA., 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.

Ohlenforst, B., Zekveld, A.A., Jansma, E.P., Wang, Y., Naylor, G., Lorends, A., Lunner, T., Kramer, SE., 2017. Effects of hearing impairment and hearing aid amplification on listening effort: a systematic review. Ear Hear. 38, 267–281.

Pichora-Fuller, M.K., Schneider, B.A., Daneman, M., 1995. How young and old adults listen to and remember speech in noise. J. Acoust. Soc. Am. 97, 593–606.

Pichora-Fuller, M.K., Kramer, S.E., Eckert, M.A., Edwards, B., Hornsby, B.W.Y., Humes, L.E., Wingfield, A., 2016. Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). Ear Hear. 37, 5S–27S.

Potts, W.B., Ramanna, L., Perry, T., Long, C.J., 2019. Improving localization and speech reception in noise for bilateral cochlear implant recipients. Trends Hear. 23, 1–18.

Rabbitt, P., 1966. Recognition: memory for words correctly heard in noise. Psychon. Sci. 6, 383–384.

Riss, D., Arnoldner, C., Baumgartner, W.D., Kaider, A., Hamzavi, JS., 2008. A new fine structure speech coding strategy: speech perception at a reduced number of channels. Otol. Neurotol. 29, 784–788.

Riss, D., Hamzavi, J.S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., Baumgartner, W.D., Gstoettner, W., Arnoldner, C., 2014. FS4, FS4-p, and FSP: a 4-month crossover study of 3 fine structure sound-coding strategies. E Hear. 35, e272–e281.

Sarampalis, A., Kalluri, S., Edwards, B., Hafter, E., 2009. Objective measures of listening effort: effects of background noise and noise reduction. J. Speech Lang. Hear. Res. 52, 1230–1240.

Schatzer, R., Krenmayr, A., Au, D.K., Kals, M., Zierhofer, C., 2010. Temporal fine structure in cochlear implants: preliminary speech perception results in cantonese-speaking implant users. Acta Otolaryngol. 130, 1031–1039.

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.

Schlueter, A., Lemke, U., Kollmeier, B., Holube, I., 2016. Normal and time-compressed speech: how does learning affect recognition thresholds in noise? Trends Hear. 20, 1–13.

Smeds, K., Wolters, F., Rung, M., 2015. Estimation of signal-to-noise ratios in realistic sound scenarios. J. Am. Acad. Audiol. 26, 183–196.

Studebaker, G.A., McDaniel, D.M., Sherbecoe, RL., 1995. Evaluating relative speech recognition performance using the proficiency factor and rationalized arcsine differences. J. Am. Acad. Audiol. 6, 173–182.

Studebaker, GA., 1985. A "rationalized" arcsine transform. J. Speech Lang. Hear. Res. 28, 455–462.

Wiggins, I.M., Seeber, BU., 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., Dorman, MF., 2018. A brief history of the cochlear implant and related treatments. Neuromodulation.

Wilson, B.S., Dorman, M.F., 2018. Chapter 99 - A Brief History of the Cochlear Implant and Related Treatments. In: Krames, E.S., Hunter Peckham, P., Rezai, A.R. (Eds.), Neuromodulation. Academic Press, pp. 1197–1207, ISBN 9780128053539. doi:10.1016/B978-0-12-805353-9.00099-1.

Zierhofer, C. (2001). Electrical nerve stimulation based on channel-specific sequences. Patent WO2001013991A1.