Multifaceted evaluation of a binaural cochlear-implant sound-processing strategy inspired by the medial olivocochlear reflex

Doctoral Thesis

María Milagros Jerónimo Fumero

Supervisor

Enrique A. López-Poveda

Instituto de Neurociencias de Castilla y León

Universidad de Salamanca

September 2020

Front page image courtesy of MED-EL GmbH

VNiVERSiDAD
DSALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL

DEPARTAMENTO DE CIRUGÍA
INSTITUTO DE NEUROCIENCIAS DE CASTILLA Y LEÓN
INSTITUTO DE INVESTIGACIÓN BIOMÉDICA DE SALAMANCA

PROF. DR. ENRIQUE A. LÓPEZ POVEDA

D. Enrique Alejandro López Poveda, con DNI 07.953.786-H, profesor titular del Departamento de Cirugía de la Universidad de Salamanca e investigador del Instituto de Neurociencias de Castilla y León,

CERTIFICA

Que la tesis doctoral titulada “**Multifaceted evaluation of a binaural cochlear-implant sound-processing strategy inspired by the medial olivocochlear reflex**” ha sido realizada por **Dña. María Milagros Jerónimo Fumero** (DNI: 78.857.715-F) bajo mi dirección.

El objetivo de la tesis es evaluar experimentalmente la audición de los usuarios de implantes cocleares con una estrategia de procesamiento binaural de sonidos inspirada en el reflejo olivococlear medial, denominada “estrategia MOC”. La tesis describe cuatro estudios dirigidos a comparar la inteligibilidad del habla en ruido, la localización de fuentes sonoras, y el esfuerzo auditivo con procesadores de sonido estándar y con diversos procesadores MOC diseñados para reflejar de forma más o menos realista el tiempo de activación del reflejo olivococlear medial natural y sus efectos sobre la compresión coclear humana. Los resultados demuestran que la estrategia MOC, con parámetros realistas, puede mejorar la localización de las fuentes de sonido y el reconocimiento del habla en ambientes ruidosos sin aumentar el esfuerzo de escucha. Además, demuestran que es posible combinar el procesamiento MOC con técnicas de codificación de audio para implantes cocleares de última generación, lo que hace que la estrategia MOC sea un enfoque prometedor para mejorar el rendimiento auditivo de los usuarios de estos dispositivos.

En conjunto, la tesis presenta una obra original, de gran calidad, que reúne los requisitos formales y el rigor científico y académico necesario para que sea defendida y optar al grado de Doctor.

Lo que firmo en Salamanca, a 14 de septiembre de 2020.

Fdo. Enrique A. López Poveda

Calle Pintor Fernando Gallego nº 1. 37007. Salamanca.
Tel.: +34 923 29 44 00 ext. #1957

audiolab.usal.es
ealopezpoveda@usal.es

VNiVERSiDAD
DSALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL

DEPARTAMENTO DE CIRUGÍA
INSTITUTO DE NEUROCIENCIAS DE CASTILLA Y LEÓN
INSTITUTO DE INVESTIGACIÓN BIOMÉDICA DE SALAMANCA

PROF. ENRIQUE A. LÓPEZ POVEDA

Solicitud de Mención de Doctorado Industrial

D. Enrique Alejandro López Poveda, con DNI 07.953.786-H, profesor de la Universidad de Salamanca, y director de la tesis doctoral titulada "**Multifaceted evaluation of a binaural cochlear-implant sound-processing strategy inspired by the medial olivocochlear reflex**" realizada por **Dña. María Milagros Jerónimo Fumero** (DNI: 78.857.715-F).

EXPONE

1. Que la tesis aborda el problema de cómo mejorar la audición de las personas que utilizan implantes cocleares mediante un procesador binaural de sonidos inspirado en el reflejo olivococlear medial, denominado "estrategia MOC". Los cuatros estudios descritos en la tesis demuestran que, comparado con el uso de procesadores de audio comerciales o de la estrategia MOC original, las implementaciones más realistas de la estrategia MOC evaluadas en la tesis mejoran el reconocimiento del habla en ambientes ruidosos y la localización de las fuentes sonoras sin requerir más esfuerzo de escucha.

2. La tesis se enmarca en el contrato de investigación (Art. 83) titulado "Procesamiento binaural de sonidos para implantes cocleares" firmado entre la empresa MED-EL GmbH y la Universidad de Salamanca, con plazo de ejecución 15/04/2013 a 31/08/2022.

3. El método de procesamiento de sonido propuesto en la tesis está patentado por la Universidad de Salamanca y este, o alguna variación de éste, podría materializarse en un nuevo tipo de implante coclear binaural, así como inspirar el diseño de nuevos audífonos e implantes auditivos. La tesis tiene, por tanto, interés industrial.

Y, por estos motivos,

SOLICITA

1. Que se otorgue a la tesis la **Mención Industrial**.

Lo que firmo en Salamanca, a 14 de septiembre de 2020.

Fdo. Enrique A. López Poveda

A MIS ABUELOS
NATALIA Y VICTORINO

ACKNOWLEDGEMENTS

Although this thesis is written in English, I have opted to write this section in Spanish because I feel more comfortable expressing my feelings in my mother tongue.

En primer lugar, me gustaría transmitir mi más sincero agradecimiento a mi director de tesis, Enrique, por haberme enseñado que, trabajando con rigor, constancia, transparencia, entusiasmo y pasión, se puede hacer buena ciencia. Gracias por ofrecerme la oportunidad de formar parte de este proyecto y por mostrarme lo bonita que es la ciencia. En definitiva, gracias por la confianza que has depositado en mí, por la paciencia, el apoyo y la ayuda que me has brindado durante estos años.

Gracias a mis compañeros del laboratorio, con los que he compartido buenos momentos y me han hecho sentir como en casa. Almudena, has sido la mejor compañera de proyecto que podía tener; no existen palabras suficientes para agradecer tu ayuda, cariño y comprensión. Peter, gracias por estar dispuesto a ayudarme siempre que lo he necesitado. Miriam, eres más que una compañera de laboratorio para mí, gracias por tu ayuda y apoyo incondicional. Alejandro, gracias por contagiarme tu buen humor, y por esos debates interminables sobre cualquier tema. Fernando y Thibaud, gracias por vuestros comentarios, sugerencias, y por el ánimo que me habéis transmitido durante este último año. Joseba, gracias por tu ayuda en lo relacionado con el ajuste de los implantes. Gracias Byanka, Helia e Inés; a pesar de que vuestra estancia en el laboratorio fue corta, habéis dejado grandes aportaciones y buenos recuerdos. Gracias a Reinhold, Josh, y Peter por vuestra ayuda con el diseño del protocolo, vuestro apoyo técnico e interés en el proyecto. Gracias a Rubén y Auxi por animar a los participantes a participar en el proyecto. Este trabajo no habría sido posible sin la ayuda de todos vosotros. MUCHAS GRACIAS.

Gracias a las personas que participaron voluntariamente en los experimentos. Ha sido un verdadero placer trabajar con vosotros. Estoy segura de que, en el futuro, todo el tiempo, trabajo y dedicación que habéis invertido en este proyecto se verán recompensados.

Pero, sobre todo, gracias a mis padres y a mis hermanos, que siempre han sido mi mayor apoyo. Mamá, gracias por ser mi motor y mi impulso; por enseñarme que, aunque algo parezca inalcanzable, con esfuerzo y sacrificio se puede conseguir. Airy, he tenido la suerte de contar con tu apoyo, comprensión y cariño. Eres mi ejemplo a seguir y el pilar fundamental de mi vida, gracias por todo lo que has hecho por mí desde que tengo conocimiento. Al resto de mi familia, gracias por vuestro apoyo y por haber confiado siempre en mí.

Álvaro, gracias por hacerme la vida tan fácil, por hacerme sentir que en Salamanca también está mi hogar, y, sobre todo, por ser mi principal apoyo durante este último año. Esta última etapa de la tesis no habría sido la misma sin tus consejos y tu ayuda incondicional.

Gracias Ana, porque junto a Miriam, sois un gran apoyo. Por estar ahí en los buenos y en los malos momentos, por lo que me hacéis reír y porque sois, en definitiva, grandes amigas.

Gracias a mis amigos tinerfeños, en especial a Elisa y Gara, por estar siempre al otro lado del teléfono transmitiéndome vuestro apoyo, pese a los meses que pasan hasta que podemos vernos.

Milagros Jerónimo Fumero, 11 de Septiembre de 2020

FINANCIAL SUPPORT

Work supported by MED-EL GmbH (Innsbruck, Austria), the Spanish Ministry of Economy and Competitiveness (Grant BFU2015-65376-P), and the European Regional Development Fund.

ABSTRACT

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 struggle understanding speech in noise or localizing sound sources with modern, bilateral CIs (BiCIs).

The MOC strategy is a binaural CI sound coding strategy inspired by the dynamic control of basilar membrane (BM) compression provided in natural hearing by the contralateral medial olivocochlear reflex (MOCR). In contrast to the standard clinical approach (STD), which involves using two independently functioning audio processors with fixed acoustic-to-electric compression, the MOC strategy dynamically couples the amount of compression applied in each ear. This can result in better speech-in-noise recognition [Lopez-Poveda et al., 2016, Ear Hear 37(3):e138-148]. Though promising, the original MOC strategy had potential drawbacks and its parameters disregarded important aspects of the natural MOCR. The aim of this thesis was to experimentally investigate the potential additional benefits of more realistic implementations of the MOC strategy for speech-in-noise recognition, sound source localization, and listening effort.

The thesis comprises four studies. The first study focused on speech-in-noise recognition. Speech reception thresholds (SRTs) for sentences presented in competition with steady-state noise were measured in unilateral and bilateral listening modes, and for multiple spatial configurations of the speech and noise sources. Speech reception thresholds were compared for stimuli processed through a STD strategy; the original MOC strategy with fast control of compression and greater inhibition at higher than at lower frequencies (MOC1); a MOC1 strategy with slower control of compression, thus closer to the time course of MOCR inhibition (MOC2); and a MOC2 strategy with greater inhibition at lower than at higher frequencies (MOC3), thus closer to the MOCR. We found that the more realistic MOC3 strategy overcame the shortcomings of the original MOC1 strategy and provided overall better speech-in-noise recognition. In addition, the MOC2 and MOC3 strategies provided a significant binaural advantage, which was not the case for the other strategies tested.

The second study focused on sound source lateralization. Bilateral CI users were asked to localize noise tokens in a virtual horizontal plane for stimuli processed through the STD, MOC1, MOC2, and MOC3 strategies. Compared to the STD strategy, the MOC1 strategy slightly improved the localization of broadband noise bursts 200 ms in duration. The MOC2 and MOC3 strategies did not improve localization because stimuli were too short to fully activate and deactivate the contralateral control of compression but could theoretically provide similar improvements for longer stimuli as the MOC1 strategy did for shorter stimuli.

The third study was aimed at investigating the potential benefits of combining MOC3 processing with a coding strategy (termed FS4) intended to preserve auditory temporal fine structure (TFS) cues in the four most apical frequency channels. Speech reception thresholds for sentences processed through the MOC3-FS4 and a standard FS4 strategy (STD-FS4) were compared in quiet, in steady-state and fluctuating noise, for various speech levels, in bilateral and unilateral listening modes, and for multiple spatial configurations of the speech and noise sources. Overall, SRTs were equal or better with the MOC3-FS4 than with STD-FS4 strategy.

The fourth study was aimed at investigating if recognizing speech in noise was as effortful with the MOC strategies as it was with the more conventional STD strategies. Word recall scores and verbal response times in a word recognition test were used as proxies for listening effort and were measured in quiet, in steady-state noise at +5 dB signal-to-noise ratio (SNR) and at the individual SRT for sentences in noise. The results showed that BiCI users experienced approximately the same effort with all sound-processing strategies.

Together, the findings show that the binaural MOC strategy, with realistic MOCR parameters, can improve sound-source localization and speech-in-noise recognition without increasing listening effort. In addition, they show that it is possible to combine MOC processing with state-of-the-art fine-structure audio coding for CIs, making the MOC strategy a promising approach to improve CI outcomes.

Keywords: cochlear implants, olivocochlear efferents, dynamic-range compression, noise, speech intelligibility, sound localization, listening effort, audio coding.

RESUMEN

Los implantes cocleares (ICs) pueden proporcionar a las personas sordas una audición eficaz mediante estimulación eléctrica directa del nervio auditivo. A pesar del progreso logrado en el diseño y el rendimiento de los ICs, los usuarios de estos dispositivos todavía tienen dificultades para comprender el habla en ambientes ruidosos o para localizar fuentes sonoras, incluso con ICs modernos y bilaterales.

La estrategia MOC es una estrategia binaural de codificación de sonido para ICs inspirada en el control dinámico de la compresión de la membrana basilar que proporciona el reflejo olivococlear medial (MOCR) contralateral en la audición natural. En contraste con el enfoque clínico estándar (STD), que implica usar dos procesadores de sonido funcionalmente independientes y con compresión acústico-eléctrica fija, la estrategia MOC vincula dinámicamente la cantidad de compresión aplicada en cada oído. Esto puede mejorar el reconocimiento de habla en ruido [Lopez-Poveda et al., 2016, Ear Hear 37(3): e138-148]. Aunque prometedora, la estrategia MOC original presenta algunos inconvenientes y sus parámetros no tienen en cuenta aspectos importantes del MOCR natural. El objetivo principal de esta tesis es evaluar experimentalmente los beneficios proporcionados por implementaciones más realistas de la estrategia MOC sobre la inteligibilidad del habla en ruido, la localización de fuentes sonoras y el esfuerzo de escucha.

La tesis consta de cuatro estudios. El primero de ellos se centró en el reconocimiento de habla en ruido. Se midieron umbrales de recepción de verbal (SRTs) para frases inmersas en ruido estacionario, en condiciones de escucha unilateral y bilateral y para múltiples configuraciones espaciales de las fuentes de habla y ruido. Se compararon los SRTs para estímulos procesados a través de la estrategia STD; la estrategia MOC original, con control rápido de la compresión y mayor inhibición en altas que en bajas frecuencias (MOC1); la estrategia MOC1 con un control más lento de la compresión, y, por lo tanto, más parecido al curso temporal de la inhibición del MOCR (MOC2); y la estrategia MOC2 con mayor inhibición en bajas que en altas frecuencias (MOC3) y, por lo tanto, más parecida al MOCR. Descubrimos que la estrategia más realista (MOC3) corrige las deficiencias de la estrategia MOC1 original y proporciona un mejor reconocimiento de habla en ruido. Además, las estrategias MOC2 y MOC3 proporcionaron una ventaja binaural significativa, algo que no ocurrió con las otras estrategias evaluadas.

El segundo estudio se centró en la lateralización de las fuentes de sonido. Se pidió a los usuarios de IC bilateral que localizaran fuentes de ruido en un plano horizontal virtual para estímulos procesados a través de las estrategias STD, MOC1, MOC2 y MOC3. En comparación con la estrategia STD, la estrategia MOC1 mejoró ligeramente la localización de ráfagas de ruido de banda ancha de 200 ms de duración. Las estrategias MOC2 y MOC3 no mejoraron la localización porque los estímulos eran demasiado cortos para activar y desactivar completamente el control contralateral de la compresión, pero en teoría podrían proporcionar beneficios similares a la estrategia MOC1, para estímulos más largos.

El tercer estudio tuvo como objetivo investigar los beneficios de combinar el procesamiento MOC3 con una estrategia de codificación de sonido (denominada FS4) destinada a preservar la estructura temporal fina del sonido en los cuatro canales de frecuencia más apicales. Los SRTs para frases procesadas a través de la estrategia MOC3-FS4 y una estrategia estándar FS4 (STD-FS4) se compararon en silencio, en ruido estacionario y en ruido fluctuante, para varios niveles de habla, en escucha bilateral y unilateral, y para múltiples configuraciones espaciales de las fuentes de habla y ruido. En general, los SRTs fueron iguales o mejores con la estrategia MOC3-FS4 que con la estrategia STD-FS4.

El cuarto estudio tuvo como objetivo investigar si el esfuerzo de reconocer el habla en ruido es menor o igual con las estrategias MOC que con las estrategias STD. El porcentaje de palabras recordadas y los tiempos de respuesta verbal en una prueba de reconocimiento de palabras se usaron como indicadores del esfuerzo, y se midieron en silencio y en ruido estacionario a +5 dB de relación señal-ruido (SNR) y en el SRT individual para frases en ruido. Los resultados mostraron que los usuarios de IC bilateral experimentaron aproximadamente el mismo esfuerzo con todas las estrategias de procesamiento de sonido.

En conjunto, los hallazgos demuestran que la estrategia binaural MOC, con parámetros realistas del MOCR natural, puede mejorar la localización de las fuentes de sonido y el reconocimiento del habla en ambientes ruidosos sin aumentar el esfuerzo de escucha. Además, demuestran que es posible combinar el procesamiento MOC con técnicas de codificación de audio para ICs de última generación, lo que hace que la estrategia MOC sea un enfoque prometedor para mejorar aún más el rendimiento auditivo de los usuarios de estos dispositivos.

Palabras clave: implante coclear, eferente olivococlear, compresión del rango dinámico, ruido, inteligibilidad del habla, localización del sonido, esfuerzo auditivo, codificación de audio.

RELATED PUBLICATIONS

Some parts of this thesis have been published in scientific journals and/or presented at international conferences, as listed below. The underlined text indicates the person who gave the conference presentation. A copy of the publications is given in Appendix 1.

PEER-REVIEWED JOURNAL PAPERS

1. Lopez-Poveda EA, Eustaquio-Martín A, **Fumero MJ**, Gorospe JM, Polo R, Gutiérrez Revilla A, Schatzer R, Nopp P, Stohl JS. (**2020**). Speech-in-noise recognition with more realistic implementations of binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. *Ear and Hearing*. https://doi.org/10.1097/AUD.0000000000000880
2. Lopez-Poveda EA, Eustaquio-Martín A, **Fumero MJ**, Stohl JS, Schatzer R, Nopp P, Wolford RD, Gorospe JM, Polo R, Gutiérrez Revilla MA, Wilson BS. (**2019**). Lateralization of virtual sound sources with a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. *Hearing Research* **379**:103-116. https://doi.org/10.1016/j.heares.2019.05.004

POSTER PRESENTATIONS

1. Lopez-Poveda EA, Eustaquio-Martín A, **Fumero MJ**, Stohl JS, Schatzer R, Nopp P, Wolford RD, Gorospe JM, Polo Lopez R, Gutierrez Revilla A, Wilson BS. Lateralization of virtual sound sources with a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. Conference on Implantable Auditory Prostheses (CIAP), 14-19 July **2019**, Granlibakken Conference Center, Lake Tahoe, CA, USA.

PODIUM PRESENTATIONS

1. Lopez-Poveda EA, Eustaquio-Martín A, **Fumero MJ**, Gorospe JM, Polo López R, Gutiérrez Revilla MA. Evaluation of alternative implementations of a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. 41 Annual Midwinter Meeting of the Association for Research in Otolaryngology, 9-14 Feb. **2018**, San Diego, CA, USA.
2. **Fumero MJ**, Eustaquio-Martín EA, Lopez-Poveda EA. Listening effort with a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. ARCHES Conference, 12-13 Nov. **2018**, Universidad de Nottingham, United Kingdom.

ABBREVIATIONS AND ACRONYMS

AGC: automatic gain control
BiCI: bilateral cochlear implant
BM: basilar membrane
CI: cochlear implant
CIS: continuous interleaved sampling
CSSS: channel-specific sampling sequence
dB: decibel
FIR: finite impulse response
FS: full scale
FSP: fine structure processing
HINT: hearing-in-noise test
HRTF: head-related transfer function
Hz: hertz
iFFM: international female fluctuating masker
IHC: inner hair cell
ILD: interaural level difference
ITD: interaural time difference
JND: just-noticeable difference
KEMAR: Knowles Electronics manikin for acoustics research
MCL: maximum comfortable loudness
MOC: medial olivocochlear
MOCR: medial olivocochlear reflex
ms: millisecond
OHC: outer hair cell
pps: pulses per second
PTA: pure-tone average
RMS: root-mean-square
s: second
s.d.: standard deviation
SNHL: sensorineural hearing loss
SNR: signal-to-noise ratio
SPL: sound pressure level
SRT: speech reception threshold
SSN: speech shaped noise
STD: standard
STOI: short-term objective intelligibility
TFS: temporal fine structure

CONTENTS

Appendix 1. Reprints of published articles

Appendix 2. Extended summary in Spanish

1.

GENERAL INTRODUCTION

1.1. MOTIVATION

Hearing is important to humans. It is essential for interpreting the surrounding environment, for communication, and for detecting dangers around. Hearing impairment, at any stage of life, can hinder communication and degrade the quality of life (Kramer et al., 2006; Nachtegaal et al., 2009; Hua et al., 2013). According to the World Health Organization (2018), over 5% of the world's population – around 466 million people – suffers from disabling hearing loss, and it is estimated that by 2050 the percentage will increase to 10%. Most people with hearing loss experience difficulties understanding speech, following conversations, and segregating concurrent sounds. This can have a negative social and emotional impact on the individual. For children, hearing loss acts like a barrier to learning spoken language, education, and social integration (Yoshinaga-Itano et al., 1998).

Sensorineural hearing loss (SNHL) is the most common type of hearing loss and is usually treated with hearing aids. Hearing aids, however, can be ineffective for people with severe to profound hearing loss and cochlear implants (CIs) must be considered instead. Cochlear implants provide a sense of hearing bypassing the damaged ear and transforming sounds into electrical pulses that directly stimulate auditory nerve fibers through electrodes implanted in the cochlea (Wouters et al., 2015).

The Food and Drug Administration of the United States of America reported that approximately 324,000 individuals worldwide had received CIs by 2012 (Sahin et al., 2017). Today, the number of CI users is probably double that figure and is expected to substantially increase in the future because of the earlier diagnosis of hearing loss, the improved CI technologies, and the reduced costs of cochlear implantation (Pisoni et al., 2017). With current CIs, most users can understand speech well in quiet environments (Baskent et al., 2016). Cochlear-implant users, however, still find it difficult to understand speech in noisy settings, localize sound sources, perceive music, or recognize speech in tonal languages, such as Chinese (Wilson, 2017). In addition, most users of bilateral CIs (BiCIs) report that hearing in noisy environments requires a high level of concentration to detect, decode, process, and comprehend speech, which increases listening effort (Perreau, 2017). This shows that bilateral CI stimulation alone is not enough to restore normal listening capabilities. The use of two independently functioning CIs with different number of frequency channels, misaligned electrodes, and/or different rates of electrical stimulation can distort or degrade binaural acoustic cues (Litovsky et al., 2012; Jones et al., 2014; Kan, 2018). The use of binaurally coordinated CI sound processors might improve the benefits from bilateral cochlear implantation.

Recently, Lopez-Poveda et al. (2016a, 2017) have shown that for some listening conditions, the intelligibility of speech in competition with other sounds can be improved by using a binaural CI sound-coding strategy termed "the MOC strategy". This strategy is inspired by (and named after) the dynamic control of basilar membrane (BM) compression provided in natural hearing by the medial olivocochlear reflex (MOCR). In contrast to the current clinical standard, which involves the use of two independently functioning audio processors with fixed compression (one per ear), the MOC strategy couples the functioning of the two sound processors and dynamically modifies the amount of compression applied in each ear. This results in an enhancement of the speech information in the ear with the better acoustic SNR (Lopez-Poveda and Eustaquio-Martín, 2018).

Lopez-Poveda et al. (2016a, 2017) showed that the MOC strategy can improve speech intelligibility for BiCI users when the target and interferer sounds are spatially separated, and for unilateral CI users when the implanted ear has the better acoustic SNR. However, the first implementation of the MOC strategy had some drawbacks: (1) it reduced the speech information in the ear with the worse acoustic SNR, which could hinder intelligibility in unilateral listening when the implanted ear has the worst acoustic SNR; and (2) the mutual inhibition between the processors decreased the overall stimulation levels and thus audibility, something that could hinder speech intelligibility in bilateral or unilateral listening modes when the two processors have identical input signals.

The original implementation and parameters of the MOC strategy disregarded aspects of the natural MOCR such as its slow time courses for activation and deactivation (Cooper and Guinan, 2003; Backus and Guinan, 2006), its greater inhibition of BM responses in apical than in basal cochlear regions (Lilaonitkul and Guinan, 2009; Aguilar et al., 2013), or that, for narrow-band MOCR elicitors, the largest inhibition occurs when the contralateral sound elicitor is one-half octave below the probe frequency (Lilaonitkul and Guinan, 2009). On the other hand, a technical evaluation of the MOC strategy predicted that the use of longer time constants for activation and deactivation of contralateral inhibition, combined with greater inhibition in the lower than in the higher frequency channels, can overcome the shortcomings of the original MOC strategy and improve the signal information in the ear with the worse acoustic SNR (Lopez-Poveda and Eustaquio-Martín, 2018). This evidence motivated the present thesis.

The overall aim of the thesis is to experimentally evaluate more realistic implementations of the MOC strategy. The specific aims are (1) to investigate the potential additional benefits of using more realistic implementations of MOC strategy for speech-in-noise recognition; (2) to compare sound source lateralization performance for different implementations of the MOC strategy (from less to more realistic); (3) to investigate the potential benefits of combining realistic MOC processing with channel-specific electrical pulse sequences designed to preserve the temporal fine structure (TFS) in speech, an approach referred to as FS4 processing and used in MED-EL CIs (e.g., Zierhofer, 2001; Riss et al., 2008; Schatzer et al., 2010); and (4) to compare listening effort during speech-in-noise recognition tasks for sounds processed with the MOC and STD strategies.

1.2. OVERVIEW OF THIS THESIS

This thesis is organized according to the objectives listed in Section 2.5. Chapter 2 sets the work framework, the hypotheses and aims. Chapter 3 describes the general methods used throughout the work, including information about the participants (CI users), the sound-processing strategies, the procedures for CI fitting and loudness-balance, and test procedures and equipment.

Chapter 4 describes a psychophysical study aimed at investigating the potential benefits of more realistic implementations of the MOC strategy for understanding speech in noise. Speech reception thresholds are reported for BiCI users and 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. Speech reception thresholds are reported for three different implementations of the MOC strategy: one with fast (MOC1) and two with slow contralateral control of compression (MOC2 and MOC3). The MOC2 strategy provided more contralateral inhibition in the higher frequency than in the lower frequency channels, while the MOC3 strategy provided more inhibition in the lower frequency than in the higher frequency channels. Speech reception thresholds for the three MOC strategies are compared to those obtained for stimuli processed through two independently functioning processors with fixed compression, an approach close the current clinical STD.

Chapter 5 describes a psychophysical study aimed at testing the hypothesis that, compared to using two independently functioning CI processors with fixed acoustic-to-electric compression (STD strategy), MOC processing enhances interaural level differences (ILDs) and improves the localization of acoustic stimuli in a virtual horizontal plane. Scores are reported for BiCI users who were asked to localize noise tokens in a virtual horizontal plane processed through the STD, MOC1, MOC2, and MOC3 strategies.

In the studies described in Chapters 4 and 5, the MOC strategy was tested 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 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. Chapter 6 assesses the possible benefits for speech-in-noise recognition of combining MOC3 with FS4 processing relative to using FS4 processing alone. These two approaches are referred to hereafter as MOC3-FS4 and STD-FS4 strategies. Speech reception thresholds for sentences in quiet and in competition with a single interferer are compared for stimuli processed through these two strategies. Speech reception thresholds are reported for steady, speech-shaped noise (SSN) and an international female fluctuating masker (iFFM), in bilateral and unilateral listening, and for multiple spatial configurations of the target and the masker sources. In addition, SRTs are reported for various speech levels to explore the potential benefits of MOC3 processing across a wider level range than was tested in the study described in Chapter 4.

Chapter 7 describes two studies aimed at comparing the effort expended by BiCI users when listening with the MOC and STD processing strategies. Effort is quantified with two different methodologies: as the score in a dual-task (word recognition and word recall) test and as the response time in a word recognition test. These studies were motivated by the fact that, despite the MOC strategy can facilitate speech intelligibility in competing noise or speech maskers in some conditions, it is not yet known to what extent this strategy can affect listening effort. It is possible that, because CI users were less familiar with the MOC strategies than with the STD or STD-FS4 strategies, they expended more effort with the MOC strategies in the intelligibility tasks. This study investigates this possibility.

Chapter 8 provides a general, integrated discussion of the main findings presented in Chapters 4 to 7, highlighting their implications, and their relationship with related studies. In addition, Chapter 8 poses new challenges and suggests possible future experiments.

1.3. ORIGINAL CONTRIBUTIONS

The main original contributions of this work are:

1. A demonstration that CI users, who lack a natural MOCR, can localize sounds and understand speech in noisy environments better when using more realistic implementations of the MOC strategy than when using conventional sound processors.
2. Evidence that more realistic implementations of the MOC strategy provide greater benefits than other, less realistic implementations evaluated to date.
3. A confirmation that CI users obtain these improvements in hearing performance without needing to expend greater listening effort.

Taken together, these findings indicate that it might be possible to improve the hearing, and thus the quality of life, of the several hundred thousand CI users (and million candidates) by using binaural audio processors inspired by the MOCR.

1.4. UNITS

All units follow the conventions of International System of Units (SI).

1.5. ANGLE CONVENTIONS

Throughout this thesis, the spatial locations the stimuli are expressed as $S_X N_Y$, where 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 azimuth angles to the right and the left of the midline, respectively.

1.6. ETHICS

Testing procedures were approved by the Western Institutional Review Board (Puyallup, WA), and by the Ethics Review Board of the University of Salamanca (Salamanca, Spain). Participants signed an informed consent before they were admitted to the studies.

2.

BACKGROUND, HYPOTHESES, AND AIMS

This chapter provides a general framework for the work presented in this thesis. First, it briefly describes the auditory pathway and explains how the healthy human auditory system works. Second, it explains the importance of binaural hearing. Third, it summarizes the impact of hearing impairment on the lives of people who suffer from it and presents a classification of hearing loss based on the site of damage in the auditory system. Fourth, it provides an overview of CIs and a brief review of conventional CI sound-coding strategies. Fifth, it reviews the advantages of bilateral over unilateral cochlear implantation, emphasizing the importance of binaural CI sound-processing strategies. Lastly, the chapter introduces listening effort and the methodologies used to measure it, focusing on single-task and dual-task paradigms.

2.1. DESCRIPTION AND FUNCTIONING OF THE HEALTHY AND PATHOLOGICAL HUMAN AUDITORY SYSTEM

2.1.1. The human auditory system

Hearing is the ability to perceive sounds by detecting vibrations through the ear. Mammals can detect and analyze sounds over a broad range of spectral frequencies and intensities. Humans can hear sounds with frequencies from 20 to 20,000 Hz and sound pressure levels (SPLs) from 0 to about 120 dB SPL (Robles and Ruggero, 2001).

To perceive sounds, the human ear transduces acoustic pressure waves into neural impulses. Sound reaches the pinna, which helps capturing environmental sounds, and travels through the ear canal, causing the tympanic membrane (or eardrum) to vibrate. This membrane forms the outer boundary of the middle ear. The vibrations of the tympanic membrane are transmitted mechanically by the ossicular chain (malleus, incus, and stapes) through the middle ear to the cochlea, a spiral-shaped bone cavity in the inner ear. The cochlea has three separate fluid compartments: the scala tympani and the scala vestibuli, which contain perilymph (similar to the body's extracellular fluid), and the scala media, which contains endolymph (similar to intracellular fluids). The scala media is separated from the scala vestibuli and the scala tympani by Reissner's membrane and the BM, respectively. The stapes lies on top of a membrane-covered opening in the cochlea called the oval window. There is a second membrane-covered opening called the round window. One of the most important functions of the middle ear is to ensure the transfer of sounds from the air to the perilymph. This transfer of sound into the cochlea depends on the difference between the sound pressure applied to the oval window and that applied to the round window. The stapes emits vibrations to the oval window, which causes pressure changes in the perilymph. When the pressure changes reach the perilymph, they give rise to a traveling wave in the BM. The movement of the oval window, produced by the movement of the stapes, results in an outward movement in the round window. This results in movement of the BM. Different positions along the BM resonate selectively to different sound frequencies, with high and low frequencies producing maximal BM vibrations at the basal and apical ends of the cochlea, respectively (von Békésy and Wever, 1960).

On top of the BM lies the organ of Corti, which contains two types of hair cells: the inner hair cells (IHCs) and the outer hair cells (OHCs) (**Fig. 2.1**). These cells have very different functions. The IHCs work as mechanoreceptors, transducing mechanical movements into neural activity. The OHCs are mechanically active receptors that determine the sensitivity and frequency selectivity characteristic of the mammalian auditory system. OHCs enhance mechanical cochlear responses to low intensity sounds, i.e., they function as a "cochlear amplifier" (Brownell, 1990; Dallos et al., 2002, 2006, 2008; Dallos, 2008). The normal functioning of the OHCs and their effect on BM vibrations help the healthy human ear to perceive sounds over a wide range of SPLs. The response of the BM in the healthy cochlea is nonlinear and compressive. This means that when the sound pressure level increases, the velocity of BM motion increases less than the increase in sound pressure level, and thereby compresses a wide range of acoustic pressure into a narrower range of BM mechanical responses (Ruggero et al., 1992; Robles and Ruggero, 2001). The amount of BM compression, however, is not constant and changes with activation of the medial olivocochlear (MOC) efferents, as described in the next section.

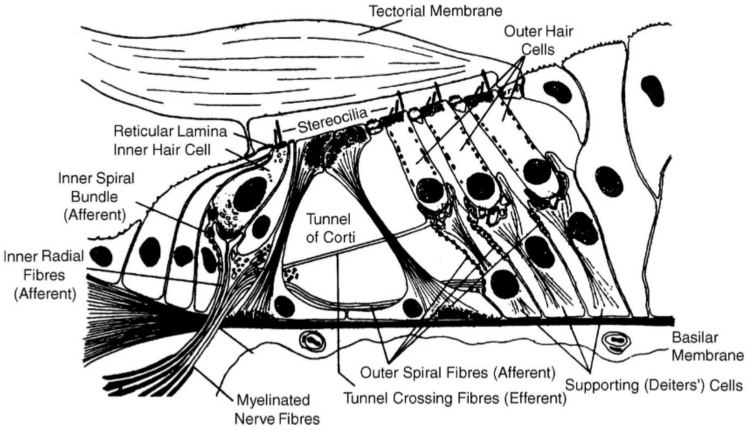

Figure 2.1. Cross section of the structure of the organ of Corti. Figure taken from Moore (2007).

2.1.2. The contralateral medial olivocochlear reflex

Our ears do not function as mere sound receptors. Instead, the central nervous auditory system can adjust and control their functioning via olivocochlear efferents (for reviews, see Cooper and Guinan, 2006; Guinan, 2006; Lopez-Poveda, 2018). The olivocochlear efferent system consists of nerve fibers which project from the brainstem to the cochlea (Warr and Guinan, 1979; Brown et al., 2010). The efferent pathway originates in the superior olivary complex, projects to the cochlea through the vestibular nerve, and terminates in the organ of Corti. The superior olivary complex is comprised of the medial nucleus of the trapezoid body, the lateral superior olive, and the medial superior olive. These two latter nuclei are the origin of fibers called the lateral olivocochlear (LOC) and the MOC efferents, respectively (Warr and Guinan, 1979; Simmons, 2002). Most MOC efferents originate in the medial superior olive and terminate directly upon OHCs, while the majority of LOC efferents originate in the lateral superior olive and terminate on the dendrites of type I auditory nerve afferent fibers, underneath IHCs (**Fig. 2.2**) (see Lopez-Poveda, 2018). Both LOC and MOC efferents contain crossed (contralateral) and uncrossed (ipsilateral) fibers although, in mammals, the majority of LOC fibers project to the ipsilateral cochlea and the majority of MOC fibers project to the contralateral cochlea. Thus, the two cochleae are connected to each other through the MOC efferents, and the functioning of each cochlea is modulated by the sounds received by the ipsilateral the contralateral ear.

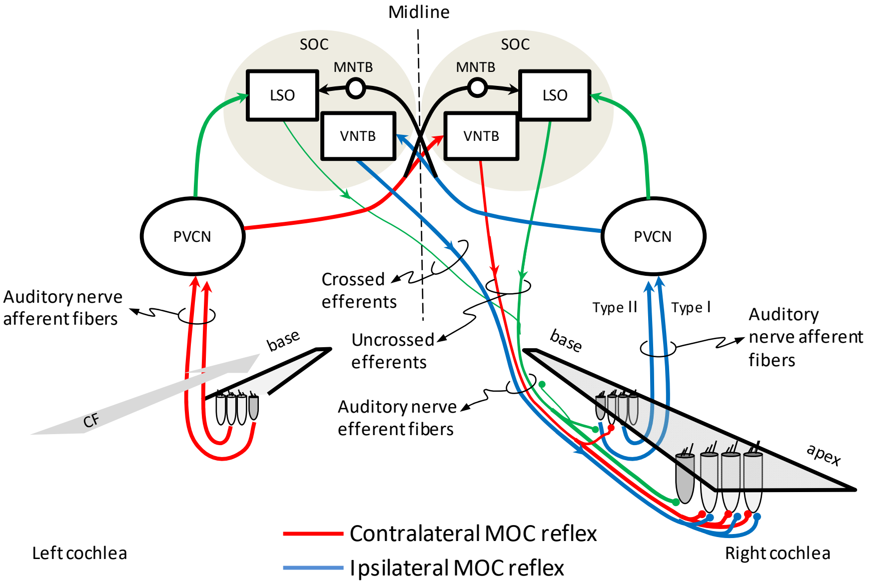

Figure 2.2. Pathways for activation of medial (MOC) and lateral olivocochlear (LOC) efferent fibers to the right cochlea. Green lines show the pathways for activation the LOC reflex. Red and blue lines illustrate the pathways for activation of contralateral and ipsilateral MOC reflexes. From Lopez-Poveda (2018).

To date, the effects of LOC efferent activation remain unclear (Guinan, 2006; Lopez-Poveda, 2018). The effects of MOC efferent activation, however, have been extensively studied. The activation of MOC efferents suppresses the electromotility of OHCs and reduces cochlear amplification to low- and mid-level sounds (Murugasu and Russell, 1996; Cooper and Guinan, 2006), which results in a linearization of BM input/output curves (Murugasu and Russell, 1996; Cooper and Guinan, 2006; Moore, 2007). For tones in continuous noise, this causes auditory nerve fibers to respond less to the noise and enhances their responses to the tones, an effect known as the “antimasking effect” of MOC efferent activation (Nieder and Nieder, 1970; Winslow and Sachs, 1988).

Medial olivocochlear efferents may be activated involuntary by ipsilateral and/or contralateral sounds (Guinan, 2006; Brown et al., 2010; Aguilar et al., 2015; Lopez-Poveda, 2018), which has given rise to the terms ipsilateral and contralateral MOCR. The MOCR is almost certainly active during natural binaural hearing, and possibly facilitates the recognition of speech in noise. In other words, similar to how it enhances the responses of auditory nerve fibers to tones in noise, the MOCR can enhance auditory nerve responses to transient speech features in noise and facilitate speech-in-noise recognition (Guinan, 2006; Kim et al., 2006; Brown et al., 2010). These “antimasking” benefits of the MOCR are probably reduced in listeners with partial OHC loss or dysfunction and absent in listeners with total OHC loss and in CI users (Lopez-Poveda et al., 2016).

2.1.3. Hearing impairment

Hearing loss is a common human sensory disease. It can hinder the development of language and degrade verbal communication and cognitive skills. In addition, it can increase the effort expended in listening, and can have a substantial socio-emotional, educational, and job-related impact. It may be acquired by adults who previously had normal hearing (post-lingual hearing loss) or it may be present at birth (pre-lingual hearing loss).

The severity of hearing loss is typically classified based on the average air-conduction hearing thresholds at 0.5, 1, 2 and 4 kHz (termed the pure-tone average or PTA) as none (or normal hearing): <15 dB HL; slight: between 15 to 25 dB HL; mild: between 26 to 40 dB HL; moderate: 41 to 70 dB HL; severe: 71 to 95 dB HL; and profound: ≥95 dB HL (Goodman, 1965; National Research Council, 2005). Hearing impairment is said to occur when the PTA is at least 35 dB higher than normal (Stevens et al., 2013).

Depending upon the site of damage in the auditory system, hearing losses can be classified as follows (Moore, 2007):

- *Conductive hearing loss* is caused by a reduction of sound transmission through the outer and/or middle ear, which produces an attenuation of sounds reaching the cochlea. It may be caused by cerumen in the ear canal, damage to the eardrum and/or the middle-ear ossicles, or by the presence of fluid in the middle ear.
- *Sensorineural hearing loss* is caused by damage to the structures or processes inside the cochlea. It can be caused by exposure to intense sounds, the use of ototoxic chemicals, infections, some allergies, or genetic factors.
- *Retrocochlear hearing loss* occurs through damage to the structures or processes beyond the cochlea, such as the auditory nerve or the central auditory system. A common cause is a benign tumor (acoustic neuroma or vestibular schwannoma) which presses the auditory nerve.

Sensorineural hearing loss is the most common type of hearing loss. Damage or destruction of the inner ear or the cochlear hair cells is the principal cause of hearing loss. If IHCs are damaged or lost, the auditory system lacks the necessary link for transforming acoustic pressure waves to neural impulses (Wouters et al., 2015). Damage or loss of OHCs results in the main characteristics of SNHL: decreased sensitivity to weak sounds, reduced cochlear frequency selectivity, suppression and compression, decreased efferent control of cochlear function and a faster growth of loudness with increasing sound level (Lopez-Poveda, 2014). Damaged hair cells can subsequently lead to degeneration of adjacent auditory neurons (Loizou, 1999).

Sensorineural hearing loss can degrade the neural representation of acoustic stimuli. This causes difficulties in daily life situations, such as loss of the ability to detect sounds, to recognize speech in noisy environments, to localize sound sources, or, more generally, to communicate with other people (Ohlenforst et al., 2017). Hearing loss increases the cognitive demands required to attend to, and to understand, an auditory message. As a result, hearing-impaired people expend extra effort to achieve successful speech understanding, something that can cause an increase in levels of mental distress and fatigue (McCoy et al., 2005). On the other hand, hearing impairment may affect the hearing-impaired person's interaction with other people, which often leads to withdrawal from social activities, rejection of visits to theatres, cinemas, lectures, etc. This reduces intellectual, cultural, and social stimulation, and leads to isolation and increased symptoms of depression (Arlinger, 2003).

It is estimated that 50% of cases with hearing loss could be prevented and most other cases could be treated effectively (World Health Organization, 2018). For the latter, most cases can be treated with hearing aids. However, when the hearing loss is severe to profound, hearing aids are not recommended because sound amplification becomes useless when the transduction of acoustic energy into neural impulses does not occur. In these cases, CIs can provide a form of treatment (Moore, 2007).

2.2. COCHLEAR IMPLANTS

The development of the CI has changed the life of profoundly deaf children and adults. This device is the most effective treatment for listeners with deafness or severe-to-profound hearing loss. Deafness is usually associated with a total loss or dysfunction of OHCs and IHCs, but the auditory nerve survives. In these cases, CIs can restore the sense of hearing via direct electric stimulation of the auditory nerve. The CI bypasses the damaged portions of the inner ear and uses direct electrical stimulation to deliver sound signals to the auditory nerve fibers and evoke a sound sensation (Moore, 2007; Loizou et al., 2009).

The early history of the CI begins with Alessandro Volta in 1790. He connected a battery with a wire that terminated with two conductive rods and placed each of the two rods within his ear canals. He experienced a sensation of sound like "boiling thick soup". This is the first report of auditory sensation elicited by electrical stimulation. In 1957, André Djourno and Charles Eyriès performed the first implant of a device for direct electrical stimulation of the auditory nerve in a totally deaf person. The patient reported auditory percepts in response to the stimuli during and after the operation, and he could hear the environmental sounds but not understand speech (Møller, 2006; Wilson and Dorman, 2018). Later, in 1972, William House and Jack Urban developed the first commercially available implants as the House-3M single channel implant. This device had a single electrode and provided electrical stimulation at a single site in the cochlea. The House-3M CI was used mostly as an aid for lipreading and provided an awareness of environmental sounds. These early devices did not facilitate the discrimination among sounds unless sounds were sufficiently different from each other in relation to their amplitude envelopes.

Today, the CI is a multielectrode device with multiple processing channels, which allows transmission of different temporal information to different locations in the cochlea. With moderns CIs, the majority of users can hold a conversation with ease in quiet environments, are able to talk on the phone, understand speech in diverse environments, and children can achieve near to normal speech and language development (Zeng, 2004; Moore, 2007; Lenarz, 2017). The modern CI is arguably the most successful neural prosthesis designed to date in that in can restore one of the human senses (Wilson, 2013, 2015, 2017). However, limitations in the device and in the peripheral auditory system, across frequency channel interactions caused by the spatial overlap of electrode stimulation, transfer limitations of the electrical information from the electrode to the auditory nerve, and cochlear abnormalities and surgical factors among others, results in a degraded signal compared to normal hearing (Baskent et al., 2016).

2.2.1. Components and functioning

The primary components of a CI device are (**Fig. 2.3**) (1) a microphone for sensing sound; (2) a sound processor to transform the microphone output into a pattern of stimuli for an implanted array of electrodes; (3) a transcutaneous transmitter of stimulus information across the skin; (4) an implanted receiver-stimulator to decode the information received from the radio-frequency signal produced by an external transmitting coil and generate stimuli using the instructions obtained from the decoded information; and (5) the electrode array (Loizou, 1999; Wilson and Dorman, 2008).

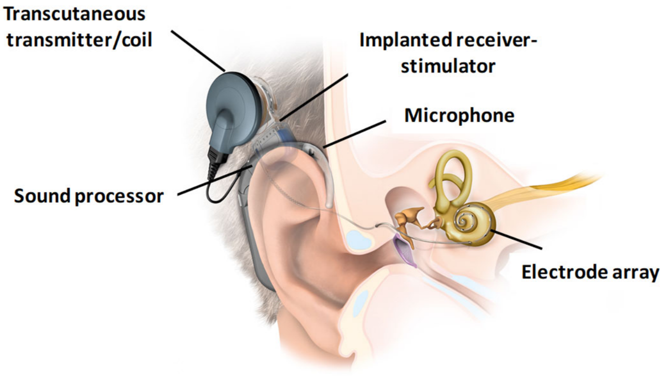

Figure 2.3. Representation of the cochlear implant system. Illustration provided by MED-EL.

The function of a CI is to bypass the eardrum, the ossicular chain, the BM, and the hair cells and directly stimulate the surviving neurons in the auditory nerve. The sound pressure variations in the environment are detected by the microphone(s) on the sound processor and converted into a digital signal. This digital signal is analyzed by an audio processor and processed through an algorithm (the sound-processing strategy) to determine the pattern of stimulation sequences that represent the sound and that should be transmitted to the implanted receiver-stimulator by a transcutaneous electromagnetic link. The main functions of the sound processor are (1) to decompose the input signal into its frequency components, and (2) to map (or compress) the wide range of acoustic pressure in the environment into the narrower range of electrical charge perceived by the CI user. The receiver-stimulator is surgically positioned posterior to the pinna and contains an electrode array (with 12-22 electrodes) inserted into the scala tympani of the cochlea. The receiver-stimulator package receives the pattern of stimulation sequences and processes the information for the stimulator to transmit it as a pattern of electrical pulses via the electrode array. The array utilizes the tonotopic arrangement of the cochlea, with low and high frequency bands stimulating the apical and the basal ends of the array, respectively (Loizou, 1999; Zeng et al., 2008; Kan and Litovsky, 2015).

The CI technology assumes that there are enough auditory nerve fibers for stimulation around the electrodes. The perceived loudness of a sound depends on the number of activated nerve fibers, which itself depends on the electric charge delivered by each electrode, which depends on the sound amplitude. Thus, loudness can be controlled by varying the electric charge. On the other hand, the perceived pitch is mostly related to the place in the cochlea that is stimulated and, to a lesser extent, to the rate of electrical stimulation (the number of electrical pulses per second or pps). Quoting Landsberger et al. (2018) "Electrodes which are placed more apically in the cochlea are reported to provide a lower pitch than electrodes which are placed more basally. Similarly, higher stimulation rates on a single electrode are reported as having a higher pitch than lower rates on the same electrode (Tong et al., 1983)".

2.2.2. Sound-processing strategies

Initial sound-coding for CIs was based on presenting the analog sound waveform picked up by the microphone to a single electrode placed in the cochlea. Later, due to the development of multichannel CIs, sound-processing strategies were divided into two categories: waveform strategies and feature-extraction strategies.

The simplest first version of multichannel CIs was based on the waveform strategy approach. In this strategy, known as the compressed analog, the spectrum of the input signals was split into 4-8 frequency bands by a bank of bandpass filters. Then, after compressing the range of sounds intensities using automatic gain control (AGC), the output of these filters was applied simultaneously to the respective electrodes. This type of processing strategy presented both spectral and temporal information. This approach, however, produced interaction between channels caused by the summation of electrical fields from individual electrodes (Loizou, 1999; Møller, 2006), which distorted the speech spectrum and degraded speech understanding. To reduce this problem, subsequent devices used pulsatile sound coding strategies. In this type of stimulation, the sound information is delivered to the electrodes using sequences of non-overlapping electrical pulses, thus minimizing channel interactions.

One of the first speech-coding approaches that used pulsatile stimulation was classified as feature extraction strategies. Vowel sounds are identified based on their formant frequencies. In these coding strategies, the fundamental (voice) (F0) and formant frequencies (F1 and F2) of speech (vowel) signals are extracted based on the assumption that these formant frequencies represent the resonance characteristics of the vocal tract during speech production. The F0 is used to control the stimulation pulse rate, while F1 and F2 determine the stimulated electrode with the assumption of a tonotopic relationship between the location and the stimulated frequency. Speech understanding with these strategies is rather poor, hence these strategies are not used in current commercial processors (Loizou, 1999; Zeng et al., 2008; Wouters et al., 2015).

In the 1990s, the CIS strategy was developed (Wilson et al., 1991). This type of processing is currently incorporated in several commercial devices (Wouters et al., 2015). The CIS strategy (**Fig. 2.4**) consists of filtering input sounds through a bank of bandpass filters with an overall bandwidth from about 100 to about 8000 Hz and a number of filters usually equal to the number of stimulation channels in the electrode array. The envelope signals extracted from the bandpass filters are compressed with a nonlinear function to map the wide dynamic range of sound in the environment (about 100 dB) into the narrow dynamic range of electrically evoked hearing. Then, the compressed envelopes are represented at the corresponding electrodes in the cochlea by trains of biphasic pulses with temporal offsets that eliminate any overlap across channels (Wilson et al., 2005; Wilson and Dorman, 2008). In other words, the pulse trains for the different channels and corresponding electrodes are interleaved in time to reduce interactions of electrical fields between electrodes to improve the specificity of place coding. A fixed stimulation rate is used (between 500-2000 pulses per second per electrode) equal for all frequency channels. In summary, the CIS strategy is based on the following principles (1) representing most of the information that can be perceived according to place, frequency, and intensity of stimulation; (2) minimizing electrode interactions; and (3) using suitable mapping functions and other aspects of processing to minimize possible distortions (Wilson, 2015).

Particularly important for the present thesis is that the CIS strategy (**Fig. 2.4**), and most other current CI sound strategies, actually include a two-stage compression approach to accommodate a broad range of acoustic pressure into the much narrower range of electrical current (Zeng, 2004). The first stage is a broadband AGC. The AGC is placed at the front-end of processing and serves to narrow the broad ranges of 'loudness' fluctuations that occur in the acoustic environment. The second compression stage is the acoustic to electric map. This map is placed at the back-end in each frequency channel of processing to map the wide dynamic range of sounds in the acoustic environment into the narrow dynamic range of electrically evoked hearing by each electrode (Wilson et al., 1991; Lopez-Poveda et al., 2016b).

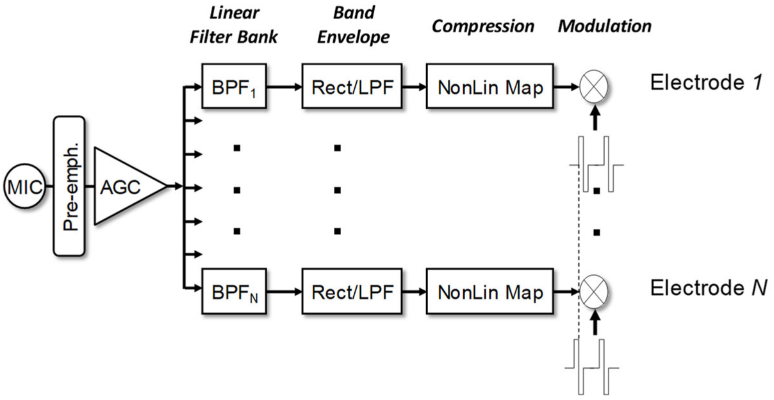

Figure 2.4. Block diagram of the elements in the standard CIS sound-processing strategy (adapted from Wilson et al., 1991). Abbreviations from left to right: MIC: microphone; AGC: automatic gain control; Pre-emph.: Pre-emphasis filter; BPF: Band-pass filter; Rect/LPF: Rectifier/Low-pass filter; NonLin Map: NonLinear amplitude mapping.

Some popular sound-processing strategies developed after CIS

Over years, several CI sound-processing strategies have been implemented and evaluated. These strategies have focused on improving the spectral and temporal representation of the input signal and on providing a better distribution of stimulation across frequency channels. Some current strategies are:

Advanced Combination Encoder (ACE). This strategy is like the CIS strategy with the exception that the outputs from a subset of filters with the highest intensities are chosen to activate electrodes in each stimulation cycle. This is referred to as an *n-of-m* approach, where *n* is the number of electrodes stimulated in each cycle and *m* is the number of frequency filters (*n*<*m*). Typically, the *n* highest-intensity filter outputs of *m* filter outputs are selected on each stimulation cycle. For example, in Cochlear$^{\text{TM}}$ CI systems, typically eight electrodes from the available 22 active electrodes are selected for stimulation at rate of 900 pps per electrode. Compared to the CIS strategy, ACE represents some speech formant peaks and changes in formant frequency over time distinctly in background noise (see Wilson et al., 1988; Mckay et al., 1991; Wilson and Dorman, 2008; Wouters et al., 2015).

Fine Structure Processing (FSP). Most CI processing strategies discard the TFS and encode speech envelopes only. The TFS, however, contains information on temporal pitch and timbre and is the primary source of interaural timing information. It also enhances speech perception in tonal languages, speech perception in noise and the quality of music perception (Wouters et al., 2015). The FSP strategy (**Fig. 2.5**) aims at improving the representation of TFS information present in the lowest frequencies of the input sound signals by delivering bursts of electrical pulses on one or several corresponding CI electrodes. This strategy uses channel specific sampling sequences (CSSS) (Zierhofer, 2001) in the lower-frequency (1-4) channels to provide envelope information as well as TFS (**Fig. 2.5B**). The timing of stimulation pulses is aligned with the zero-crossing of the band-pass filter outputs in the apical low-frequency channels (TFS is presented for frequencies up to 350-500 Hz), so TFS information is transmitted at low frequencies. MED-EL has launched two processing strategies called FS4 and FS4-p, which present TFS information for frequencies up to 750-950 Hz (see Hochmair et al., 2006; Wilson and Dorman, 2008; Lorens et al., 2010; Schatzer et al., 2010; Wouters et al., 2015).

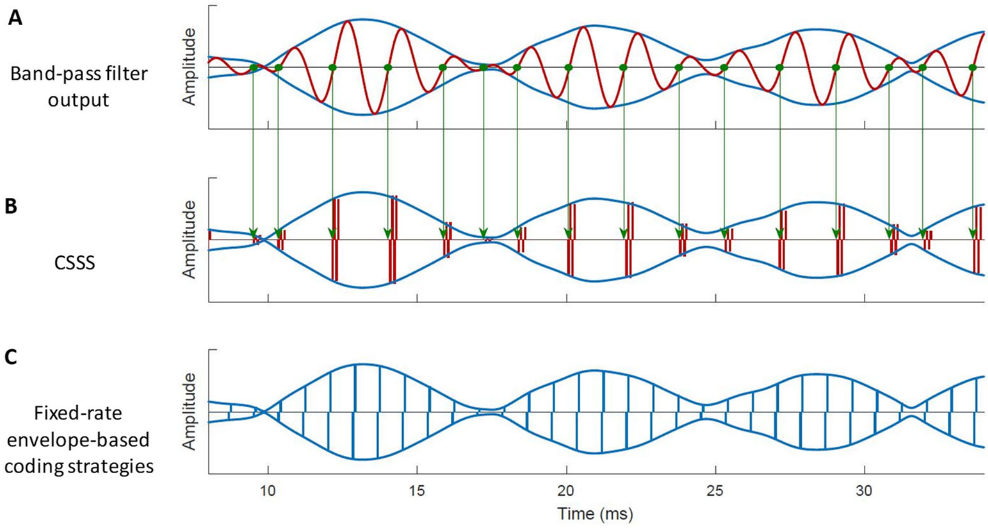

Figure 2.5. Difference between the pulse trains of the CIS strategy (**C**) with those of MED-EL's FSP strategy (**B**). Figure provided by MED-EL GmbH. CSSS: Channel-specific sampling sequences.

MP3000. The MP3000 strategy is based on the ACE strategy but uses a psychoacoustic masking model derived from findings for normal-hearing listeners to select the sites of stimulation in the implant. The idea is that it should not be necessary to code sounds in parts of the spectrum that are masked. This would improve sound perception based on more perceptually relevant selection of channels. This processing strategy recruits *n-of-m* spectral components but takes *n* components with the highest levels relative to the calculated masked threshold. Therefore, this approach selects the most important spectral components in any given input audio signal and can lead to a more precise representation of the spectrum (see Nogueira et al., 2005; Wilson and Dorman, 2008; Buechner et al., 2011; Wouters et al., 2015).

HiRes120. This strategy aims at improving the transmission of spectral information. It makes use of virtual channels to increase the number of stimulation sites within the cochlea, thus increasing the spatial resolution of stimulation and perception with CIs. It is incorporated in Advanced Bionics commercial devices. HiRes120 uses 15 electrode pairs (if all 16 intra-cochlear electrodes are active) and current steering creates eight additional stimulation sites between each of these electrode pairs, amounting to 120 potentials sites of stimulation for the array. The stimulation rate is typically about 2000 pps per electrode (see Koch et al., 2004; Koch et al., 2007; Wouters et al., 2015).

Main peak interleaved sampling (MPIS). This strategy was developed by the Neurelec-MXM company and is the basic strategy employed in Oticon Medical CI systems. It is a conceptually different strategy from the previous ones. It uses current pulses of identical amplitude and controls the loudness by varying the duration (thus the charge) of each electrical pulse. In this processing strategy, the speech processor adaptively selects the number of channels with the highest amplitude or ‘maxima’ depending on the amplitude characteristics of the incoming sounds. For each maximum, the electrical pulse amplitude (electrical current) remains constant while the electrical pulse duration is dynamically adapted for each patient. The number of maxima may vary between 1 and 20 in a total of 20 electrodes. Stimulation rate may be varied from 150 to 1000 Hz. The use of high stimulation rates should allow MPIS strategy to combine detailed spectral and temporal signal information (Di Lella et al., 2010).

Five manufacturers of CIs are on the international market and each uses slightly different strategies: Cochlear (ACE, MP3000), MED-EL (CIS, FSP); Advanced Bionics (HiRes120), Oticon Medical (MPIS), and Nurotron (CIS, ACE, and virtual channels) (Zeng et al., 2008; Zeng et al., 2015; Wouters et al., 2015). This listing is not comprehensive. The development of different CI processing strategies has led to significant improvements in hearing over the years. Conventional sound-coding strategies, however, are monaural and CI users continue to have limitations understanding speech in noisy environments, especially when they cannot take advantage of binaural hearing (Sarant et al., 2014).

2.3. BINAURAL HEARING

The human auditory system is binaural, that is, it is specialized in processing sound received at two ears. Binaural hearing is the faculty to take advantage from comparisons of the acoustic signals received at the two ears.

2.3.1. The importance of binaural hearing for spatial localization

Sound source localization is an important aspect of everyday life. To localize sound sources in the horizontal plane, normal-hearing people rely on acoustic cues arising from differences in arrival times of a sound wave at the two ears (also known as inter-aural time differences or ITDs), and on acoustic cues arising from differences in the level of stimuli at the two ears (known as inter-aural level differences or ILDs). Inter-aural time difference and ILD cues are not equally useful at all frequencies. Low-frequency sounds (<1500 Hz) are localized using ITDs, whereas high-frequency sounds (>1500 Hz) are localized using ILDs. Inter-aural time and level differences depend not only on sound frequencies but also on the size of the head and on the sound source’s azimuth angle (Akeroyd, 2006; Avan et al., 2015).

2.3.2. The importance of binaural hearing for speech-in-noise recognition

For people with normal hearing, binaural hearing is essential for better understanding speech in background noise (Hawley et al., 2004; Laszig et al., 2004; Brown and Balkany, 2007). The advantages related with binaural hearing encompass three effects: binaural redundancy, head shadow effect, and binaural 'squelch'.

Binaural redundancy or binaural summation

Binaural redundancy refers to the central auditory system´s ability to benefit from duplicate representations of the same signal at the two ears, i.e., from diotic stimulation. For a perfectly symmetrical head, it occurs when speech and noise originate from the same spatial location in the median sagittal plane. With diotic stimulation, there is redundancy in the information and an enhanced sensitivity to changes in intensity and frequency that contribute to improved detection and speech recognition. The doubling of perceptual loudness is perceived as an increase of 3 dB in quiet environments compared with monaural listening (Brown and Balkany, 2007).

The head shadow effect

The head shadow effect occurs in everyday listening conditions when speech and noise sources are spatially separated. When the speech and noise sources are placed at different locations, the presence of the head produces a sound diffraction pattern that leads to different SNRs in the two ears. The SNR is higher in the ear further away from the noise source and smaller in the ear closer to the noise source. There may be a difference around 15 dB SNR between the two ears, and this difference decreases as the speech and noise sources get closer. The head shadow effect is disadvantageous for people with unilateral hearing loss when the target sound is located on the impaired side, and this disadvantage increases in presence of background noise (Akeroyd, 2006; Avan et al., 2015).

Binaural 'squelch'

The binaural 'squelch' is defined as the increase in speech recognition when listening with two ears compared to listening with the ear that has the better SNR (i.e., the acoustically better ear) alone. For spatially separated noise and target sound sources, the acoustically better ear is that closer to the target source and the acoustically worse ear is that closer to the noise source. Binaural squelch results from the brain's ability to take advantage of the different time and intensity cues of the auditory inputs that reach each ear when the speech and noise are spatially separated (Laszig et al., 2004; Litovsky et al., 2006; Avan et al., 2015).

2.3.3. Are two cochlear implants better than one?

Numerous studies have shown that most unilateral CI users have no difficulties to develop speech, language, or communication skills, or to understand speech in quiet settings (Svirsky et al., 2004; Peterson et al., 2010). However, unilateral CI users do not benefit from the abovementioned binaural advantages and often find it difficult to maintain a conversation in noisy settings or to locate sound sources (Müller et al., 2002; Litovsky et al., 2004; Schleich et al., 2004). An attempted solution to this problem has been bilateral cochlear implantation.

The main motivation for bilateral implantation is to restore spatial hearing abilities. Bilateral implantation, however, also enhances auditory sensitivity and speech-in-noise intelligibility, and subjectively improves sound quality and the quality of life (Schleich et al., 2004; Litovsky et al., 2004, 2006, 2012; Brown and Balkany, 2007; Kan, 2018; Smulders et al., 2016). Regarding speech intelligibility, the most robust binaural advantage from bilateral cochlear implantation is the head-shadow effect (Gantz et al., 2002; Müller et al., 2002; Schleich et al., 2004; Eapen et al., 2009). When using two rather than one CIs, head shadow alone improves speech reception thresholds (SRTs) in noise by 4-7 dB SNR, a benefit similar to that provided by head shadow to normal-hearing listeners (around 8 dB SNR) (MacKeith and Coles, 1971; Bronkhorst and Plomp, 1988). The next most robust advantage is binaural summation. In BiCI users, it varies between 1.5 and 2.9 dB (Gantz et al., 2002; Müller et al., 2002; Schleich et al., 2004; Eapen et al., 2009). These values are smaller than those for normal-hearing listeners (3-6 dB) (MacKeith and Coles, 1971; Bronkhorst and Plomp, 1988). The third most robust binaural advantage is the binaural squelch. The magnitude of this effect is around 3 dB for normal-hearing listeners (Bronkhorst and Plomp, 1988), and only a small number of BiCI users show squelch (Müller et al., 2002; Laszig et al., 2004; Litovsky et al., 2004; Schleich et al., 2004). Buss et al. (2008) and Eapen et al. (2009), however, showed that binaural squelch increases significantly beyond the first-year post-implantation. In relation to spatial location abilities, BiCI users have better sensitivity to ILDs than to ITDs than unilateral CI users (Grantham et al., 2007; Dorman et al., 2014).

In summary, despite two CIs are better than one for speech intelligibility and spatial localization, BiCI users still do not perform as well as normal-hearing people do in similar tasks (Litovsky et al., 2004; Loizou et al., 2009; Kan, 2018). This may be because there is technical (software and hardware) limitations in what CIs can achieve (Wilson and Dorman, 2008) but also because the two CIs worn by BiCI users function independently from each other, thus as two monaural systems rather than as a binaural system.

2.3.4. Binaural processing strategies

Current CI sound-processing strategies limit the benefits of binaural hearing, partly because they do not fully preserve ITDs and ILDs (Brown and Balkany, 2007). In other words, for users of two CIs, each of the two devices operates as a monaural system, which degrades and/or distorts binaural cues (van Hoesel, 2004; Litovsky et al., 2012).

Sound-processing strategies have been developed to improve the representation of ITD cues for BiCI users. One such strategy was the peak derived timing (PDT) strategy (van Hoesel and Tyler, 2003). This strategy aims at preserving fine-structure ITD information by synchronizing stimulation pulses from the CI with amplitude peaks in the fine structure of the signals in the different channels of the filter bank. The authors demonstrated that the PDT strategy produces moderate ITD sensitivity (of the order of 100 µs), but this sensitivity was deteriorated when stimulation rates increased above a few hundred Hertz. Another strategy developed to improve the perception of ITDs in bimodal stimulation[1] is the modulation enhancement strategy (MEnS) (Francart et al., 2014).

---

[1] Bimodal stimulation refers to the use of a CI in one ear with a hearing aid in the other ear.

This strategy enhances the perception of ITDs for bimodal listeners by modulating the electric stimulation signal synchronously with modulations in the acoustic signal presented to the non-implanted ear. The MEnS strategy improved ITD thresholds and sound source lateralization compared to the ACE strategy. These studies (and several others not cited here) show that the use of bilaterally synchronized devices can improve the perception of ITDs in experimental laboratory tests. Inter-aural time difference sensitivity with these devices, however, continues to be worse than that of normal-hearing people.

Similarly, several strategies have been proposed to enhance ILDs. Francart et al. (2011) developed an algorithm that enhances ILDs for bimodal CI users. They showed that ILD enhancement improved sound source localization performance by 4° to 10°, relative to a mean absolute error of 28° without ILD enhancement. Another method for enhancing ILDs at low frequencies was proposed by Moore et al. (2016). This method was tested using simulated hearing-aid processing in bilateral hearing-aid users. The algorithm did not improve speech intelligibility but improved the localization of speech sources by a few degrees. Brown (2018) proposed a sound-processing strategy with the aim to provide BiCI users with larger than normal ILD cues. This strategy improved localization performance. Moreover, two of the six subjects that participated in the study achieved localization performance levels typical of those observed in normal-hearing listeners. Dieudonné and Francart (2018) developed a method to enhance head-shadow ILDs in the low frequencies using a fixed beamformer with contralateral subtraction in each ear. The method was tested on normal-hearing listeners simulating bimodal stimulation. In the localization task, the angle error improved from 50.5° to 26.8°. Speech reception thresholds in noise improved by 15.7 dB SNR when the noise was presented from the CI side, 7.6 dB SNR when the noise was presented from the hearing-aid side and were not affected when noise was presented from all directions. Another novel listening strategy for improving speech-in-noise recognition in BiCI users inspired by the *better ear* phenomenon has been developed by Kan (2018). This strategy combines the knowledge of a better ear with a signal-processing algorithm that separates the target talker from the noise and delivers the target to the better ear while the remaining sound is presented to the contralateral ear. Speech recognition performance was evaluated using a virtual auditory space in BiCI users. Speech reception thresholds in noise improved by 4.4 dB.

Lopez-Poveda et al. (2015, 2016b) developed a binaural CI sound-processing strategy that reinstates some of the effects of the contralateral MOCR on cochlear compression to CI users: “the MOC strategy”. This strategy is the focus of the present thesis and so its functioning and benefits are described in detail in the next section.

2.3.5. A binaural CI sound-processing strategy inspired by the contralateral MOCR: The MOC strategy

In natural hearing, cochlear mechanical compression is dynamically adjusted via the contralateral MOCR (Guinan, 2006). These adjustments possibly facilitate understanding speech in noisy environments but are not currently available to the users of CIs because the electrical stimulation delivered with the CI bypasses OHCs (the site of action of the MOC efferent fibers) and is thus independent from the MOCR. Lopez-Poveda et al. (2016a, 2016b) have shown that the use of a binaural sound strategy inspired by the MOCR facilitates the intelligibility of speech in competition with spatially separated maskers both in unilateral and bilateral listening conditions. This strategy can also improve the segregation of concurrent, spatially separated speech and noise sources (Lopez-Poveda et al., 2017).

The MOC Strategy

Current CI sound processors roughly reproduce the basic signal processing performed by a healthy ear. Unfortunately, these processors ignore a fundamental aspect for communication in noisy environments: the involuntary (or reflexive) control over the operation of each ear exerted by sound received through the contralateral ear. Unlike what happens in the normal auditory system, where BM compression is dynamic by the action of the MOCR, compression in conventional CIs is fixed for all sound inputs (Lopez-Poveda, 2015). Because the electrical stimulation delivered with the CI is independent from cochlear mechanical processes, the adjustment of compression provided by the MOCR in natural hearing is unavailable to CI users (Wilson et al., 2005). Perhaps this contributes to the greater difficulties experienced by CI users communicating in noisy environments. However, MOCR effects can be reinstated using dynamic compression (Lopez-Poveda et al., 2016a).

The MOC strategy is a binaural CI sound-coding strategy that uses dynamic (time-varying), binaurally coupled back-end compression inspired by the inhibitory effect of the MOCR on BM responses (Lopez-Poveda, 2015). Currently, bilateral CI users wear a pair of audio processors that function independently from each other and with fixed compression in each frequency channel of processing. We will refer to this approach as the standard (STD) strategy. In contrast with the STD approach, in the MOC strategy, the compression in each frequency channel of each processor varies dynamically in time depending upon the output level from the corresponding frequency channel in the contralateral processor (**Fig. 2.6**). The coupling is such that the greater the level at the output of every frequency channel of processing in an audio processor, the more linear the back-end compression in the corresponding frequency channel of the contralateral audio processor (Lopez-Poveda et al., 2016a, 2016b; Lopez-Poveda and Eustaquio-Martin, 2018). In other words, the greater the output level in the contralateral ear, the greater the inhibition relative to the STD strategy and the smaller the stimulus amplitude in the ipsilateral ear (**Fig. 2.7**).

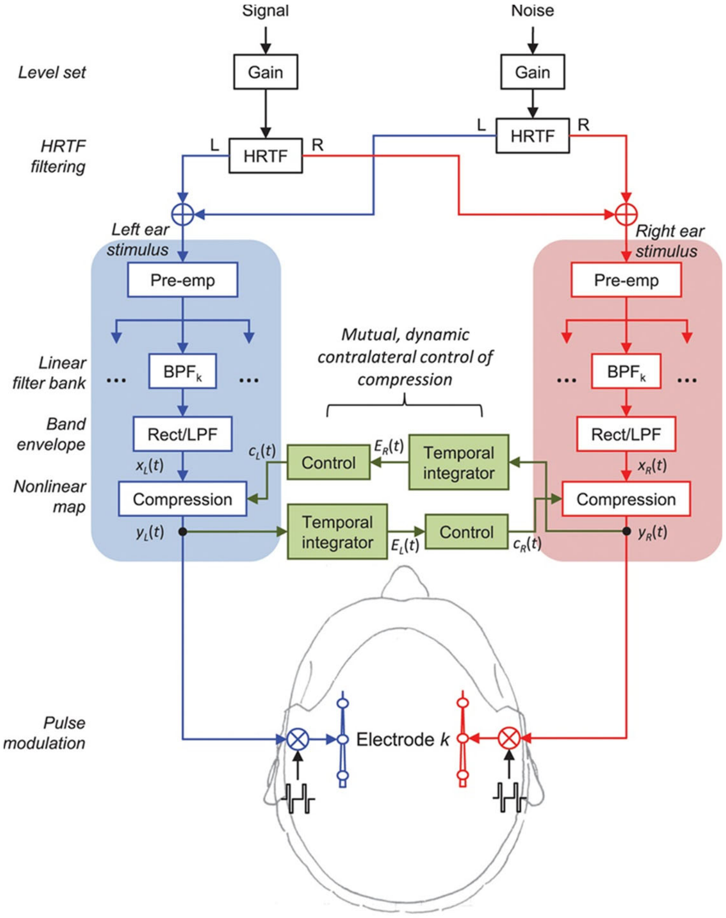

Figure 2.6. Signal processing in the STD and MOC strategies (from Lopez-Poveda et al., 2016b). Note that STD and MOC processors were identical except that MOC processors included contralateral control of back-end compression (processing is only shown for the *k*th channel). The diagram also illustrates how monophonic speech and noise signals were filtered through appropriate head-related transfer functions (HRTF) to simulate free-field stimuli during the experimental tests. See the main text for further details.

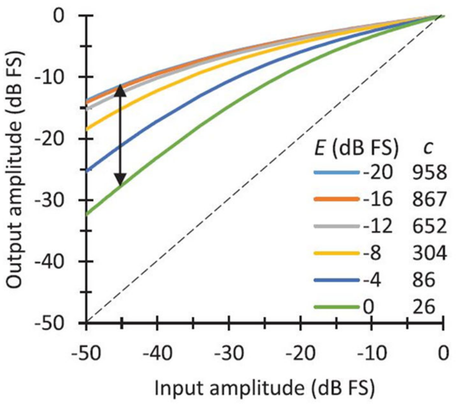

Figure 2.7. Range of instantaneous compression functions (Eq. 3.1) for six different values of the contralateral output level (*E*) linearly distributed from −20 to 0 dB FS and corresponding values of the parameter *c*, as shown in the inset. The double-headed vertical arrow illustrates the amount of inhibition for an input level of −45 dB FS. dB FS means dB re unity. FS indicates full scale; MOC, medial olivocochlear reflex. Figure taken from Lopez-Poveda et al., (2016b).

To date, the MOC strategy has been implemented and tested using short (2 ms) time constants for the activation and deactivation of the contralateral inhibition. Compared to using two independently functioning processors with fixed compression (STD strategy), the MOC strategy enhanced the speech information in the ear with the better acoustic SNR. The perceptual tests conducted so far with CI users have demonstrated that compared to the STD strategy, the MOC strategy facilitates the recognition of speech in competition with both steady-state noise (**Fig. 2.8A,C)** and a single-talker masker (**Fig. 2.8B,D**) for bilateral and unilateral CI users when the target and masker sound sources were spatially separated, and when the implanted ear had the better acoustic SNR. (Note that the MOC strategy is binaural, i.e., it requires capturing and processing the signals at the two ears, but the processed signals may be used as appropriate to stimulate the implanted ear of unilateral CI users).

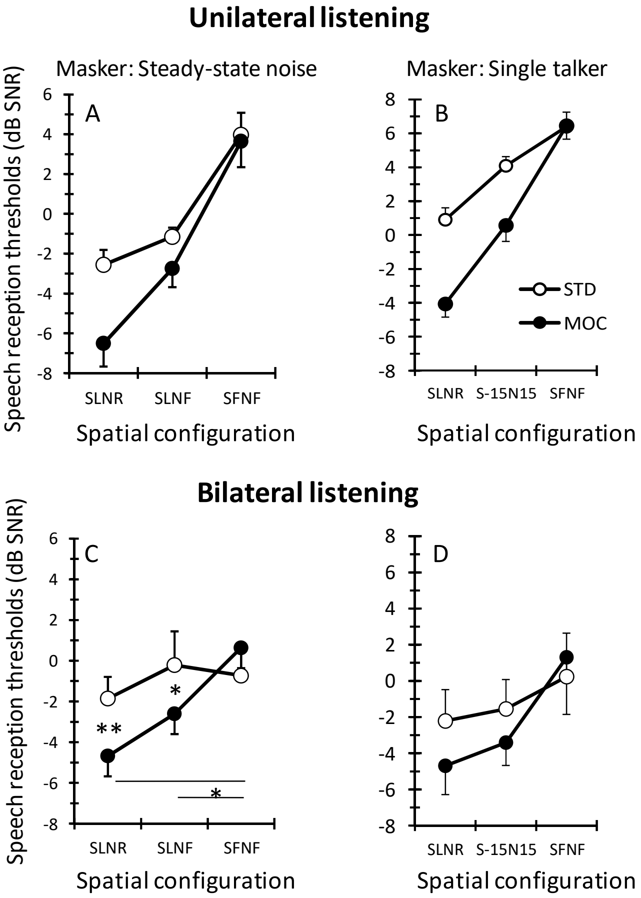

Figure 2.8. Mean SRTs measured with the STD and MOC strategies for speech in competition with steady-state noise and a single talker masker. **A, B.** Mean scores in unilateral listening. **C, D.** Mean scores in bilateral listening. Each symbol is for a different processing strategy. Left panels are for steady-state noise masker. Right panels are for a single-talker masker. SLNR indicates speech to the left ear with noise to the right ear; SLNF, speech to the left ear with noise in front; SFNF, speech and noise in front; S−15N15, speech 15° to the left with noise 15° to the right from the midline; MOC, medial olivocochlear reflex; STD, standard strategy. Data from Lopez-Poveda et al. (2016b) and Lopez-Poveda et al. (2017).

However, this implementation of the MOC strategy had some disadvantages: (1) it reduced the speech information in the ear with the worse acoustic SNR, which could hinder intelligibility in unilateral listening when the implanted ear has the worst acoustic SNR; and (2) the mutual inhibition between the pair of processors decreased the overall stimulation levels and thus audibility, something that could hinder intelligibility in bilateral or unilateral listening when the two CIs (or processors) have identical input signals.

The original implementation and parameters of the MOC strategy disregarded aspects of the natural MOC reflex including its slow time course for activation and deactivation (Cooper and Guinan, 2003; Backus and Guinan, 2006) or that it causes greater inhibition of BM responses in apical than in basal cochlear regions (Lilaonitkul and Guinan, 2009; Aguilar et al., 2013). A technical evaluation of the MOC strategy using the short-term objective intelligibility (STOI)[2] (Taal et al., 2011) predicted that MOC processing could provide even wider benefits with more realistic implementations of natural MOC effects (Lopez-Poveda and Eustaquio-Martín, 2018). It predicted that the use of longer time constants for activation and deactivation of contralateral inhibition, combined with comparatively greater inhibition in the lower than in the higher frequency channels can overcome the shortcomings of MOC processing and even improve the signal information in the ear with the worse acoustic SNR (Lopez-Poveda and Eustaquio-Martín, 2018). The principal aim of this thesis is to confirm these predictions by experimentally evaluating more realistic implementations of the MOC strategy with BiCI users in different auditory tasks.

2.4. LISTENING EFFORT

Listening effort is frequently defined as “the mental exertion required to attend to, and understand, an auditory message” (McGarrigle et al., 2014). A recent consensus paper proposed the Framework for Understanding Effortful Listening (FUEL) to address many of the complexities that go into concepts of spoken communication, including listening effort. This framework defines listening effort as “the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task, with listening effort applying more specifically when the task involves listening” (Pichora-Fuller et al., 2016). Note that this definition accentuates the difference between the demand of a given listening situation and the effort deliberately exerted by a listener.

When the quality of the auditory input is reduced because of an impaired auditory system or because of adverse acoustical environments, listeners usually deploy extra cognitive resources to understand an auditory message or localize a sound source (Hicks and Tharpe, 2002; Winn, 2016). For instance, individuals with hearing impairment often report difficulties understanding speech in unfavorable listening conditions and devote more concentration and attention when listening in those unfavorable conditions than in quiet (Gagné et al., 2017). In addition, they usually must allocate more cognitive resources to listen to and understand a conversation than do people with normal hearing. This increase in cognitive resources has been referred to as an increase in listening effort (Hicks and Tharpe, 2002; Fraser et al., 2009), and it can be associated with fatigue (Hornsby, 2013). Thus, effortful listening can lead to negative consequences for the listener and their active participation in society (Hua et al., 2014). For this reason, the inclusion in clinical settings of listening

---

[2] The STOI (Taal et al., 2010, 2011) is an average linear correlation coefficient computed over time frames between a time-frequency representation of unprocessed clean speech and processed noisy speech. That is, the STOI is the average correlation between an ideal template (e.g. the unprocessed speech in quiet) and the processed speech in noise (Lopez-Poveda and Eustaquio-Martín, 2018).

effort measures as dimensions of hearing impairment can be valuable for carrying out and report on the effectiveness of the interventions. The factors that may contribute to increased listening effort include, but are not limited to, age (Larsby et al., 2005; Tun et al., 2009; Gosselin and Gagne, 2011), hearing impairment (Tun et al., 2009; Desjardins and Doherty, 2013; Picou et al., 2013; Ohlenforst et al., 2017), device factors (Sarampalis et al., 2009; Pals et al., 2012; Hornsby, 2013; Desjardins, 2015; Picou et al., 2017; Ohlenforst et al., 2017), motivation (Picou and Ricketts, 2014; Eckert et al., 2016), fatigue (Tharpe et al., 2006; Hornsby, 2013; Eckert et al., 2016; Gustafson et al., 2018), and attention and working memory capacity (Kahneman, 1973; Rudner et al., 2012; Picou et al., 2013; DiGiovanni et al., 2017; Koelewijn et al., 2017). The influence of multiple factors on the experience of listening effort suggests that effort might be a multidimensional process (Peelle, 2018; Alhanbali et al., 2019).

There are several benefits of assessing listening effort. For example, it is useful to determine intervention strategies, to inform and discuss stressful situations for patients, or to provide evidence that intervention is needed. Furthermore, listening effort can also be included as a dimension when developing new processing strategies to improve the listening experience of the patient. For these reasons, over the last years, hearing researchers and clinicians are becoming more interested in the concept of listening effort and the usefulness of measuring it as part of clinical evaluation (Gosselin and Gagné, 2010; McGarrigle et al., 2014). However, there is no standardized procedure to measure listening effort (Pichora-Fuller et al., 2016).

2.4.1. How is listening effort assessed?

A variety of subjective and objective paradigms have been used to understand, estimate and document the effect of listening effort in different acoustic situations. However, there is no clear rationale for choosing among the different options, and often these measures of listening effort do not correlate with each other (Alhanbali et al., 2019). This raises questions about the validity and sensitivity of the different measures. The lack of agreement between measures also suggests that listening effort might be a multidimensional phenomenon with different measures evaluating independent aspects of the same process (Alhanbali et al., 2019). The most common approaches to assess listening effort include subjective, physiological, and behavioral methods (**Table 2.1**) (McGarrigle et al., 2014; Pichora-Fuller et al., 2016).

Table 2.1. Examples of the different measures of listening effort.

Listening Effort Measures	Tasks/Examples
Subjective measures	Self-reported listening effort experience: closed-set questionnaires or rating scales.
Physiological measures: changes in the central nervous system	Brain activity measures: evoked-response potentials, electroencephalography, and functional magnetic resonance imaging.
Physiological measures: changes in the autonomic nervous system	Sympathetic and parasympathetic nervous system responses: pupil, cardiac, skin, and hormonal responses.
Behavioral measures (single-task)	Response time to verbal, tactile or/and visual stimuli.
Behavioral measures (dual-task)	Perform two tasks concurrently. The primary task is the experimental listening task of interest. The secondary task is used as a competing task, which usually involves performance on either a memory task or response to a stimulus.

Subjective measures. Several studies have used rating scales and questionnaires to subjectively estimate the individual amount of perceived effort during a listening experience. Currently, if listening effort were measured in clinical environments, it would be likely done with self-reports or rating scales designed to measure the amount of perceived effort, acceptance, benefit and satisfaction with the hearing device provided for treatment (Gosselin and Gagné, 2010). These self-report measures tend to be closed-set questionnaires that are often related to daily life experiences and offer a closed set of response opportunities [e.g. the Speech, Spatial and Qualities of Hearing scale (SSQ; Gatehouse and Noble, 2004)]. These measures have a good face validity, are quick and easy, provide first-hand information, and can be administered without specialized test equipment. However, one limitation of this measure is the lack of correlation between the objective and the subjective measures of listening effort. Zekveld et al. (2011) reported no correlation between changes on pupil size and self-reported listening effort when presenting sentences to participants at different levels of background noise. The findings suggest that pupillometry and self-reported measures assess independent and possibly different aspects of listening effort. Another limitation is that self-report measures may be affected by the way effort is interpreted by the participants. Thus, the individual may be influenced by their current state of mind, rather than the state induced by a given testing condition (McGarrigle et al., 2014; Pichora-Fuller et al., 2016). For example, a participant in a relatively positive state of mind might underestimate the extent of their effort exerted to the listening task and vice versa. Larsby et al. (2005) suggest that older people tend to underestimate their perceived effort because in their study, older participants did not report a greater degree of perceived effort than the younger participants did, despite a demonstrated group difference in the objective measures. On the other hand, self-report measures of listening effort do not explain the physiologic process underlying this phenomenon (Bess and Hornsby, 2014). Thus, the influence of these limitations on the subjective measures of listening effort might have contributed to the lack of correlation between self-report methodologies and other physiological or behavioral measures of effort (Mackersie and Cones, 2011; Mackersie et al., 2015).

Physiological measures. The assumption underlying physiological indices of effort is that any cognitively and/or perceptual challenging task will reveal physiological changes that occur during task performance (Zekveld et al., 2011). These changes can be attributed to an increase in listening effort (Pichora-Fuller et al., 2016). In these paradigms, the participant is performing a listening task and measurements of bodily fluctuations (e.g. pupil diameter or skin conductance) are simultaneously recorded. These measures of listening effort fall into two main categories: measures of brain activity and measures of the autonomic nervous system (McGarrigle et al., 2014). The most common techniques for measuring neural activity to infer listening effort are magnetic encephalography, evoked-response potentials, electroencephalography, and functional magnetic resonance imaging. Measures of the autonomic nervous system may tap sympathetic and parasympathetic nervous system responses, which can be influenced by changes in task demand. These autonomic responses can be measured using pupil, cardiac, skin conductance, and hormonal responses (Pichora-Fuller et al., 2016).

Magnetic encephalography and evoked-response potentials measurements have been used to study time-locked neural activity evoked by the presentation and the response to stimuli. Electroencephalography shows the response to acoustic stimuli measured by electrodes on the scalp and provides precise markers of mental processing during a task (Bernarding et al., 2013; Miles et al., 2017). Functional magnetic resonance imaging has been used to assess the role of effortful listening by observing metabolic consequences of neuronal activity and changes in blood oxygenation level during an auditory task. This technique has been used to assess the effect of CI simulations on speech processing on effortful listening (Wild et al., 2012). On the other hand, pupil diameter has been considered a measurable index of cognitive processing load, for example, in relation to changes in attention and perception (Kramer et al., 1997; Zekveld et al., 2010; 2011). This technique assumes that when speech processing is cognitively demanding, pupil dilation will be observed. There is evidence that pupil size systematically changes in relation to task-evoked load and mental effort during a task (Zekveld et al., 2010; Koelewijn et al., 2012). Pupillometry has been previously used to assess how hearing impairment (Kramer et al., 1997; Zekveld et al., 2011), speech intelligibility (Zekveld et al., 2010), different types of background noise (Koelewijn et al., 2012), dynamic environment like a cocktail party (Koelewijn et al., 2015), and auditory attention (Koelewijn et al., 2017) influence and affect listening effort. An advantage of this technique is that it provides an indication of changes in cognitive demand during perception tasks. For this reason, pupillometry may become a complementary measure to subjective ratings of effort in clinical settings (Miles et al., 2017). Regarding skin conductance, a response is assessed by measuring the conduction of electrical current through the skin. Mackersie and Cones (2011) observed an increase in skin conductance and heart rate when task demand increased. On the other hand, Hicks and Tharpe (2002) measured cortisol levels extracted for saliva samples and observed that this measure is associated with cognitive demands and fatigue as a response to stressors in children with and without hearing loss. However, if the goal is to find a listening-effort measure suitable for clinical purposes, physiological measures would require expensive equipment and the procedures could sometimes be uncomfortable.

Behavioral measures. Listening effort can also be assessed objectively using behavioral measures. These measures can be classified into single-task and multi-tasking paradigms (McGarrigle et al., 2014). These measures are described below.

2.4.2. The single-task paradigm

This paradigm is based on the assumption that challenging listening tasks result in increased response time as a result of increased cognitive demands, and thus increased listening effort (Gatehouse and Gordon, 1990; Houben et al., 2013; Gustafson et al., 2014). This methodology is simple: the participant is instructed to listen to an auditory stimulus (e.g., words or sentences) and repeat them out loud. Verbal responses are recorded, and verbal response time is measured as the time lapse between the offset of the auditory stimulus and the onset of the verbal response (Meister et al., 2018). Gatehouse and Gordon (1990) were the first to use this behavioral measure to evaluate the benefit obtained from hearing aid use. Response times were measured in unaided and aided conditions. The task was the detection of pure tones and recognition of single words and sentences in SSN. They found that the response times of hearing-impaired listeners were shorter with the use of hearing aids than without hearing aids, suggesting a decrease in listening effort on speech intelligibility for the aided condition. Houben et al. (2013) measured the influence of noise on the response time for both an intelligibility task and an arithmetic task in normal-hearing listeners. The arithmetic task was used to increase the cognitive demands associated with task performance. There was a significant effect of SNR on the performance of both tasks, with slower response times in the more adverse SNRs. Also, they showed that response times were longer in the arithmetic task compared to the identification task. Altogether, these findings suggest that the differences in response times might be related to listening effort and might be used to evaluate hearing-aid signal processing. On the other hand, Pals et al. (2015) compared two behavioral response time measures of listening effort that can be combined within a clinical speech test: verbal response times to auditory stimuli and response times to a visual task in a dual-task paradigm. Both measurements showed longer response times in the presence of noise. Due to the ease of implementation, verbal response time seems to be a useful effort measure to complement clinical speech tests.

The use of response times as a measure of listening effort may be advantageous in clinical settings compared to other measures, since response times can be relatively simple to obtain (Houben et al., 2013; Pals et al., 2015). The principal limitation of using the response time as a measure of effort is that it is not a "pure process" of listening effort; that is, multiple aspects can influence the speed of processing including age (Pichora-Fuller et al., 2016). A second limitation is that response time measures may be sometimes insensitive 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. In addition, it is also possible that increased effort to maintain task performance may result in shorter response times (Bess and Hornsby, 2014).

2.4.3. The dual-task paradigm

The concept of dual-task paradigms is based on the theory that the total processing resources a person has available to perform tasks are limited in capacity and processing speed to respond to a stimulus (Kahneman, 1973). That is, humans have a limited capacity for processing information and any task that requires capacity (space) in the processing mechanism will interfere with other task that also requires space. This type of experimental paradigm provides a form of ecological validity to the experimental procedure because dual tasking is often required when processing speech in real-life situations (Gagné et al., 2017). However, the effort measured in the laboratory is likely different from effort experienced in natural environments due to differences in the duration of the tasks and the complexity of natural listening situations (e.g. noisy environments, attend to multiple-talkers, etc.) (Pichora-Fuller and Singh, 2006).

The classic dual-task paradigm requires a participant to perform two tasks concurrently. One task is the *primary* task that is usually the experimental listening task of interest and involves listening in different acoustic conditions (e.g., speech recognition task). The *secondary* task is used as a competing task and varies across studies, but usually involves performance on either a memory or response task. Based on the assumption of the dual-task paradigm, an increase in cognitive load for the primary task will deteriorate performance in the secondary task as more cognitive resources are diverted to support the execution of the primary task. Thus, a deterioration in secondary task performance is inferred to reflect an increase in listening effort (Gosselin and Gagne, 2011).

It can be argued that the cognitive domains on which the dual-task paradigm is based fall into three categories that are interrelated: attention, working memory, and processing speed (Pichora-Fuller et al., 2016). Attention is important and is involved in the ability to select and maintain information during one activity (selective attention) or multiple activities (divided attention) (Pichora-Fuller et al., 2016). Related to attention, different secondary tasks have been used to accompany the primary tasks: visual task (Hornsby, 2013; Picou et al., 2013; Picou and Ricketts, 2014), tactile pattern recognition task (Fraser et al., 2009; Gosselin and Gagne, 2011), and driving a vehicle simulator (Hicks and Tharpe, 2002; Fraser et al., 2009; Gosselin and Gagne, 2011). On the other hand, working memory capacity is limited and can be assigned to processing and storing information during the performance of complex activities. There is a variety of secondary tasks used within the working memory dual-task paradigms and these assume that if more capacity is allocated to listening (primary task), less spare capacity will remain available for storing information (secondary task). Recall is a popular task to measure listening effort in dual-task paradigms (Sarampalis et al., 2009; Hornsby, 2013). Recall dual-task paradigms have been used in the evaluations of the effects of hearing loss and the performance of CI and hearing-aid users in listening effort (Lunner, 2003; Hughes and Galvin, 2013; Perreau et al., 2017).

The dual-task paradigm has several limitations. First, it can be affected by individual differences in aspects such as task engagement and motivation (Alhanbali et al., 2019). Reviews of studies using dual-task paradigms have demonstrated that behavioral measures suffer from imprecision and results are difficult to compare across studies (Ohlenforst et al., 2017). Second, different measures of listening effort are not usually correlated with each other (McGarrigle et al., 2014; Alhanbali et al., 2019). For example, performance on the dual task often does not correlate with self-reported listening effort (Gosselin and Gagne, 2011; Hornsby, 2013). The lack of relation between the dual-task methodology and self-reported measures suggest that these two measures may be evaluating independent aspects of listening effort. Third, 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 identify whether participants use all their cognitive capacity or not. Fourth, it is not possible to know with certainty if the participant always prioritizes the performance of the primary task. Fifth, the dual-task paradigm may not be the method of choice for use in a clinical setting because performing two tasks simultaneously can be difficult to do and the procedure difficult to explain to some populations, such as children or the elderly (Alhanbali et al., 2019).

In summary, there is no reliable measure of listening effort. It would be important to reach a consensus on which methods are optimal or at least appropriate to measure listening effort in order to guide researchers to define research objectives and design future studies, as well as to assist clinicians in improving their practice related to the listening effort measures (Ohlenforst et al., 2017).

2.4.4. Listening effort in CI users

Several studies have shown that CI users typically spend more effort than do normal-hearing people performing the same auditory task (Hughes and Galvin, 2013; Perreau, 2017). One reason may be that CI users suffer from poorer than normal speech-in-noise intelligibility. The MOC strategy can improve the recognition of speech in noise for CI users (Lopez-Poveda et al., 2016a, 2016b, 2017), but it remains to be assessed how much effort it takes for CI users to listen with the MOC strategy. On the one hand, CI users may need to exert less effort when listening with the MOC than with the STD strategy because the MOC strategy facilitates the recognition of speech in noise. On the other hand, however, CI users may need to exert less effort when listening with the STD than with the MOC strategy because the STD strategy is closer to the audio processing strategy in their clinical devices.

One aim of this thesis is to compare listening effort with the various implementations of the MOC strategy and with a STD strategy during speech-in-noise recognition tasks. Listening effort for BiCI users was assessed using a dual-task paradigm, but a single-task paradigm (response time) was also used in a subset of tests conditions (see **Chapter 7**).

2.5. AIMS AND HYPOTHESES

The main aim of this thesis is to experimentally evaluate various implementations of the binaural MOC strategy designed to reflect more or less realistically the control of BM compression exerted by the contralateral MOCR in natural, acoustic hearing. To achieve this general aim, the following specific objectives were established:

1. To investigate the potential benefits of using more realistic implementations of the MOC strategy for speech-in-noise recognition in unilateral and bilateral listening modes. This includes investigating the binaural advantages provided by MOC processing.
2. To experimentally verify if virtual sound source localization in the horizontal plane is better with the MOC than with STD sound-processing strategies.
3. To investigate the potential benefits of combining realistic MOC processing (termed MOC3 strategy) with FS4 processing relative to using FS4 processing alone. The benefits will be investigated for speech-in-noise recognition, with different types of maskers and in unilateral and bilateral listening conditions.
4. To compare listening effort in speech-in-noise recognition tasks for sounds processed with the MOC and STD strategies.

The overall hypothesis is that the benefits of MOC processing (relative to the STD strategy) will be greater with more realistic implementations of natural MOCR effects.

The specific hypotheses are:

1. In the MOC strategy, the use of longer time constants for activation and deactivation of contralateral inhibition, combined with comparatively greater inhibition in the lower than in the higher frequency channels, overcome the shortcomings of the initial implementation of the MOC strategy and even improve the signal information in the ear with the worse acoustic SNR. Therefore, more realistic implementations of the MOC strategy will produce better performance in speech-in-noise recognition tasks in unilateral and bilateral listening conditions and for various spatial configurations of the target and masker stimuli.
2. Compared to the STD strategy, the MOC strategies enhance ILD cues and improve the localization of acoustic stimuli in a virtual horizontal plane. Sound lateralization performance will be better with more realistic implementations of the MOC strategy.
3. The MOC3 strategy in combination with the FS4 sound-processing produce better SRTs in noise that FS4 processing alone.
4. Because the MOC strategy facilitates the recognition of speech in noise, listening in noise with this strategy requires the same or less effort than listening with the STD strategy.

3.

GENERAL METHODS

Chapters 4 to 7 report four different studies, each one aimed at addressing one of the four aims listed in Section 2.5. This chapter describes the processing strategies, participants and general methods used across studies. It should be noted, however, that not all processing strategies and participants were tested in all studies. For this reason, and to facilitate the reading, the specific methods for each study are described in corresponding chapter, maintaining participant codes and processing strategy naming across chapters.

3.1. PARTICIPANTS

Twenty users of bilateral and two bimodal users of MED-EL CIs participated in the studies (**Tables 3.1** and **3.2**). Three of them were tested at the MED-EL US Laboratory (North Carolina, USA), and 19 of them were tested at the University of Salamanca (Salamanca, Spain). All participants reported to perform very well with their implants. Participants were volunteers and not paid for their services. The participants tested at the University of Salamanca were native speakers of Castilian Spanish, while those tested at the MED-EL US Laboratory were native speakers of American English. Tests were distributed in two experimental protocols. **Table 3.1** and **Table 3.2** show data for the participants in the first and second protocol, respectively. Only one person participated in the two protocols (identified as SA014 in the first protocol and SA022 in the second protocol).

Table 3.1. Data for the participants in the first protocol. Participants whose IDs start with ME and SA were tested in North Carolina and Salamanca, respectively. Better ear indicates the better ear as reported by the participant. Ab: antibiotics; B: behavioral; *c:* compression parameter value in the participants clinical CIs; Ch: cholesteatoma; F: female; Ge: genetic; HA: hearing aid; He: hereditary; Inf: infections; L: left; M: male; MCL: maximum comfortable loudness; Mg: meningitis; n/a: not applicable; Nn: neurinoma; Os: otosclerosis; pps: pulses per second per electrode; R: right; Syn: syndromic; Un: unknown; Vol.: volume.

ID	Sex	Age (years)	Etiology	Time of implant use (months)		Num. Electrodes used for testing		Pulse rate (pps)		Better ear	*c* value in the clinical devices		THR (%MCL)		Vol. (%)	
First protocol				**L**	**R**	**L**	**R**	**L**	**R**		**L**	**R**	**L**	**R**	**L**	**R**
ME115	M	81	Un/He	47	47	9	9	1587.3	1587.3	R	1000	1000	0	0	100	100
ME131	M	54	Un/He	30	32	11	11	1578.9	1823.7	L	1000	1000	0	0	100	100
ME132	M	43	Un	62	62	9	9	1587.3	1587.3	R	1000	1000	B	B	92	100
SA004	F	35	Ge	22	13	11	11	1550	1567	R	500	500	10	10	125	120
SA005	M	44	Mg	119	103	11	11	1600	1504	R	500	500	0	0	110	100
SA006	F	48	Ge	HA	125	n/a	11	n/a	1653	R	n/a	1000	n/a	5	n/a	130
SA007	M	49	Ge	HA	125	n/a	11	n/a	1617	R	n/a	1000	n/a	15	n/a	130
SA008	M	16	Un	13	129	10	10	1818	1020	R	500	500	10	5	130	100
SA009	M	15	Ge	105	148	10	10	1818	1538	R	500	500	0	10	125	130
SA010	M	16	Un	140	172	10	10	1695	1099	R	500	500	10	0	130	130
SA011	F	44	Un/Ab	22	135	10	10	1754	1734	L	500	600	5	5	110	120
SA012	F	7	Ge	76	65	12	12	1515	1485	L	500	500	5	5	90	100
SA013	M	8	Ge	83	83	12	12	1485	1515	R	500	500	10	10	110	110
SA014	M	48	Mg	175	190	9	9	1846	1143	L	900	500	5	5	100	120
SA015	F	35	Mg	147	19	11	11	1405	1653	L	1000	500	5	5	110	110
SA016	F	74	Un/He	150	119	10	10	1493	1478	L	500	500	10	10	110	110

Table 3.2. As **Table 3.1** but for the participants of the second protocol. Note that SA022 was identified as SA014 in **Table 3.1**.

ID	**Sex**	**Age (years)**	**Etiology**	**Time of implant use (months)**		**Electrodes Active/ Used for testing**		**Pulse rate (pps)**		**Better ear**	***c* value in the clinical devices**		**THR (%MCL)**		**Vol (%) STD MOC**	
Second protocol				**L**	**R**	**L**	**R**	**L**	**R**		**L**	**R**	**L**	**R**	**L**	**R**
SA021	F	40	Inf	29	175 12	1-12 1-12	1-12 1-12	1399	1210	L	500	500	10	10	85 95	75 85
SA022/ SA014	M	49	Mg	185	200	1-9 1-9	1-7,9-11 1-7,9-10	1500	1322	L	900	500	10	10	90 95	75 80
SA023	F	68	Os	211	184	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
SA024	M	62	Nn/Ch	119	96	1-9 1-9	1-12 1-9	1268	1449	L	500	500	10	10	90 90	80 80
SA025	F	16	Un	169	180	1-9 1-8	1-8 1-8	1266	1293	R	500	500	10	10	90 95	90 95
SA026	M	19	Un	52	69	1-11 1-11	1-11 1-11	1258	1382	R	500	500	10	10	95 100	90 95
SA027	F	19	Syn	44	19	1-11 1-11	1-12 1-11	1302	1240	R	500	500	10	10	85 90	90 100

3.2. SOUND-PROCESSING STRATEGIES

3.2.1. *STD strategies*

Two standard strategies were tested, one that disregarded TFS cues (STD) and one intended to preserve TFS cues (STD-FS4). The two standard strategies involved using two functionally independent sound processors (one per ear), each one with fixed back-end compression.

The **STD strategy** was based on the CIS strategy (Wilson et al., 1991). The two processors (see **Fig. 2.4**) in the pair included a high-pass pre-emphasis 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 cut-off 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 (described below); and CIS of the compressed envelopes with biphasic electrical pulses. Note that this strategy was implemented without AGC.

The **STD-FS4 strategy** included a bank of MED-EL's proprietary finite impulse response (FIR) bandpass filters with a modified logarithmic distribution between 70 and 8500 Hz; envelope extraction via Hilbert transform; a channel-specific gain to the input signal to the compression function (this gain replaced the high-pass filter employed in the STD strategy); a logarithmic compression function; and CIS of compressed envelopes with biphasic electrical pulses using the FS4 approach, i.e., using CSSSs in the four most apical channels and fixed-rate stimulation sequences in the remaining channels. This strategy was implemented with linked AGC, meaning that the AGC functions at the two ears applied identical gain equal the minimum gain across the ears.

The number of filters in the filter banks were identical to the minimum number of active electrodes between the left and right implants (**Tables 3.1 and 3.2**), and equal between the left- and right-ear processors.

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

$$y = \frac{ln(1+c\cdot x)}{ln(1+c)} \qquad , \qquad (3.1)$$

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 and STD-FS4 strategies, 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 Salamanca participants in their clinical devices (shown in **Tables 3.1 and 3.2**).

Note that the STD and STD-FS4 strategies were the most similar to those employed by the participants in their clinical devices, except for the use of a linked AGC in the STD-FS4 strategy. We, however, used research implementations of the CIS and FS4 strategies that could be slightly different from the implementations of the corresponding strategies in the clinical audio processors.

3.2.2. MOC strategies

The MOC processors were as the STD processors except that the compression parameter value (c in Eq. 3.1) in every frequency channel of processing varied dynamically depending upon the time-weighted output level from the corresponding frequency channel in the contralateral processor (**Fig. 2.6**). 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. 2.7**). Specifically, c varied between approximately 30 and 1000 for contralateral output levels of 0 and –20 dB FS (where 0 dB FS means 0 dB re unity), respectively, as in the previously published experimental studies of the MOC strategy (see Lopez-Poveda et al., 2016b, 2017).

Inspired by the exponential time-course of activation and deactivation of the MOCR (Backus and Guinan, 2006), in the MOC strategy, the instantaneous output level from the contralateral processor was calculated as the root-mean-square (RMS) output amplitude integrated over a preceding exponentially decaying time window with two time constants (τ_a and τ_b, $\tau_a \leq \tau_b$).

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 pseudo-logarithmic distribution of band-pass filter center frequencies, high-frequency channels had larger bandwidths than low-frequency channels. As a result, for broadband signals, the output level could have been greater for the higher frequency than for the lower frequency channels. This, together with the high-pass pre-emphasis filter, could make the output level and thus the contralateral inhibition, to be greater for higher frequency than for lower frequency channels.

To better control the amount of contralateral inhibition 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, i.e., c depended on E' rather than E, where E' was calculated as follows:

$$E' = E \cdot \sqrt{\frac{BW_{ref}}{BW}} \quad , \qquad (3.2)$$

where BW is the channel bandwidth, and BW_{ref} is the bandwidth of a reference frequency channel. Unless otherwise stated, the BW_{ref} was approximately equal to the bandwidth of median channel (the actual normalization channel was #7, #6, #5 and #5 for participants with 12, 11, 10 and 9 active channels, respectively). This produced effectively greater inhibition in the lower than in the higher frequency channels, as illustrated in Figure A1 of Lopez-Poveda and Eustaquio-Martín (2018).

Specific implementations of the MOC strategy

The principal aim of this thesis was to evaluate various implementations of the MOC strategy designed to reflect more or less realistically the inhibitory characteristics of the natural MOCR. Four different implementations were tested:

- **MOC1.** This was the MOC strategy as implemented and tested originally (Lopez-Poveda et al., 2016b, 2017), with fast time constants ($\tau_a = \tau_b = 2$ ms) and with greater inhibition in the higher than in the lower frequency channels. Bandwidth normalization (Eq. 3.2) was not applied.
- **MOC2.** This was a MOC1 strategy (i.e., without bandwidth normalization) with longer time constants $\tau_a = 2$ ms, $\tau_b = 300$ ms, thus closer to the slower time course of activation and deactivation of the natural contralateral MOC reflex (Backus and Guinan, 2006).
- **MOC3.** This was a MOC2 strategy with bandwidth normalization (Eq. 3.2) 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 and Guinan, 2009b).
- **MOC3-FS4** was a MOC3 strategy with binaurally linked AGC and with FS4 processing to preserve the TFS in the four most-apical frequency channels.

Importantly, the MOC1, MOC2 and MOC3 strategies had the STD strategy as the reference, while the MOC3-FS4 strategy had the STD-FS4 strategy as the reference.

3.3. FITTING AND LOUDNESS LEVEL BALANCE

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 individually measured values, or to 0, 5, or 10 percent of MCL values (Boyd, 2006) (**Tables 3.1 and 3.2**). Processor volumes were set using the STD or STD-FS4 strategies (see below) to ensure that sounds at the two ears were perceived as comfortable and equally loud, and that a sentence filtered with the HRTF 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. Thresholds, MCL levels, and processor volumes remained constant for each participant across test conditions.

When the STD strategy was used as the reference (protocol 1), processor volumes remained identical for the STD and MOC1, MOC2 and MOC3 strategies. This was to ensure that the contralateral inhibition produced the corresponding 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 normal-hearing listeners (Aguilar et al., 2015; Kawase et al., 2003; Smith et al., 2000). By contrast, when the STD-FS4 strategy was the reference (protocol 2), volumes were fitted separately for the STD-FS4 and MOC3-FS strategies (**Table 3.2**) in an attempt to compensate for the possible reduction in loudness associated with MOC3 processing.

3.4. 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. Speech reception thresholds were measured using fixed-level speech and varying the noise level adaptively using a one-down, one-up procedure. For each SRT measurement, thirty 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 ten 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 initial SNR was 20 dB. 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 standard deviation 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 SRT were measured in this way for each condition and the mean was of three measures was regarded as the final SRT. For the two children, only one SRT was measured per condition.

When appropriate, SRTs were measured also in quiet using a similar adaptative procedure but varying the speech level rather than the SNR. The initial speech level was such that the participants could understand a full sentence, typically –20 dB FS.

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.

3.5. SOUND SOURCE LOCALIZATION PROCEDURE

Sound source localization performance was measured using a virtual acoustic space. Participants were presented with eight noise tokens for each one of 11 azimuthal angles in the frontal horizontal plane (88 noise tokens in total). The 88 noise tokens were presented in random order. During the presentation of the stimuli, participants sat in front of a computer screen that displayed a top view of a human head with an array of speakers in front of the head (**Fig. 3.1**). For each stimulus presentation, the subject was instructed to judge the azimuthal position of the sound source by clicking on the corresponding speaker in the computer screen. The click of a response triggered the processing of a freshly generated noise stimulus through the corresponding strategy, and the presentation of the processed stimulus to the participant. In North Carolina, the response screen displayed 11 speakers spaced every 15° over an azimuth range from –75° to 75°. In other words, the range of possible responses was equal to the range of actual azimuth locations. In Salamanca, the screen displayed an array of 37 speakers spaced every 5° over an azimuth range from –90° to 90° (**Fig. 3.1**), even though stimuli were presented at azimuths from –75° to 75° every 15°. The latter approach was used to increase the angle error at chance performance.

In North Carolina, tests were conducted in two blocks of four measurements per angle (i.e., two blocks of 44 presentations) per strategy. (Participants ME115 and ME132 performed the tests in the order STD, STD, MOC1, and MOC1, while ME131 performed the tests in the order MOC1, STD, STD, and MOC1). In Salamanca, tests were conducted in one block of eight measurements per angle per strategy (88 presentations in total) and the different processing strategies were tested in random order. Additional precautions were taken to minimize potential learning effects that might have biased scores across strategies. First, participants were encouraged to train themselves on the task by clicking on any of 37 speakers evenly spaced every 5° over an azimuth range from –90° to 90° and listening to the corresponding stimulus; that is, during training, participants could hear stimuli at all those azimuthal locations while for testing, stimuli were presented at a subset of locations. Training was provided independently for each processing strategy and for as long as each participant deemed necessary. Second, during the measurements, feedback was not given to participants on the correctness of their responses. Third, participants did not know which processing strategy they were training on or being tested with. Fourth, in Salamanca, the full protocol (four strategies and 88 stimulus presentations per processing strategy) was administered twice for all participants and the results of the first round were regarded as practice and discarded from further analysis.

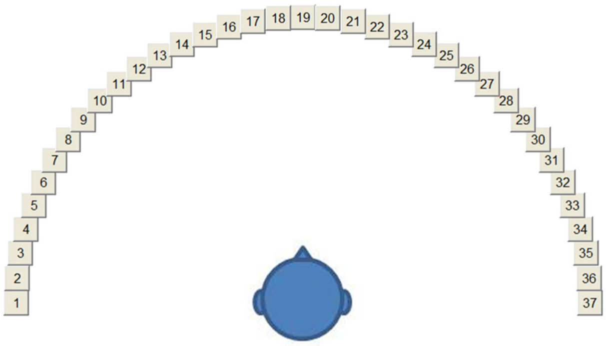

Figure 3.1. The response window for the localization experiment used in the Salamanca laboratory.

3.6. LISTENING EFFORT PROCEDURES

Listening effort was assessed using two objective methodologies.

3.6.1. Dual-Task: 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), the standard for clinical testing in Spain[3]. Words uttered by a male

---

[3] 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.

talker were presented in quiet or in competition with SSN. The experimenter controlled the presentation of the stimuli. 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 played, 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 assume that the proportion of recalled words informs about the amount of effort spent by the listener in the recognition task. Three measurements were made per participant and per condition, and the mean was taken as the final score.

3.6.2. Single-task paradigm: Verbal response time

Response time was used as an alternative assessment of listening effort during the dual-task measurement. 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. Sounds indicating hesitation or thinking were not regarded as a response in computing response times. Response times were measured for each of the ten words that were presented in each test condition, regardless of whether the word was recognized or not. Three measurements were made per participant and test condition, and the mean was taken as the final score. Longer response times were interpreted to reflect greater effort in the word recognition task.

3.7. VIRTUAL ACOUSTICS

All stimuli were directly delivered to the participant's implant via a research interface (see below). For the first protocol, free-field listening was simulated by filtering monophonic (speech) recordings through diffuse-field equalized head-related transfer functions (HRTFs) for a Knowles Electronics Manikin for Acoustic Research (KEMAR), and for speakers 1 m away from the center of the manikin's head (Gardner and Martin, 1995). For the second protocol, spatial configurations were achieved by convolving monophonic recordings with Opus 3 (o3) HRTFs provided by MED-EL. In all cases, the speech and noise sources were at eye level (i.e., their elevation angle was 0°). 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, however, results are reported as if the better-ear was the right ear for all participants.

Unless otherwise stated, unilateral listening tests involved using the self-reported better ear.

3.8. EQUIPMENT

The MATLAB software environment (versions R2014a and 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).

Because all stimuli were directly delivered to the participant's implant, sound insulation was not necessary. For this reason, during the measurements, participants were seated in a regular room and the experimenter was typically sitting in front of the participant controlling the experimental software and scoring the participant's responses if appropriate.

3.9. DOUBLE-BLIND APPROACH

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

3.10. STATISTICAL ANALYSES

All statistical analyses were conducted using IBM SPSS Statistics version 23.

In **Chapters 4** and **6** 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, 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. All tests were two-tailed, and a result was regarded as statistically significant when $p \leq 0.05$.

When data did not conform to a normal (Gaussian) distribution, non-parametric Friedman's test for related samples and related groups was applied to test for the effects of processing strategy on SRTs. For completeness, Friedman's two-way analysis of variance (ANOVA) by ranks was also applied to test for the effects of processing strategy and spatial configuration on SRTs. SPSS does not apply correction for multiple comparisons in connection with Friedman's two-way ANOVA by ranks. For this reason, post-hoc analyses were conducted using Wilcoxon signed-rank tests without Bonferroni corrections. Nonetheless, Bonferroni corrections were applied *ad hoc* by dividing the criterion p value for statistical significance by the number of made comparisons (given by the product of strategies times spatial configurations). For example, for N comparisons and $p \leq 0.05$, a result would be statistically significant if the obtained p value is smaller than $p \leq 0.05/N$.

In the sound-source localization study (**Chapter 5**), Kolmogorov-Smirnov tests (with Lilliefors correction) were used to test if the distributions of angle error and Pearson's correlation coefficient between actual and response azimuth were normal. When this happened, parametric repeated-measures analyses of the variance (RMANOVA) and/or paired Student's t tests were used to test for the statistical significance of processing strategy on angle error or correlation scores.

4.

SPEECH INTELLIGIBILITY WITH MORE REALISTIC IMPLEMENTATIONS OF THE MOC STRATEGY[4]

4.1. INTRODUCTION

As reviewed in the General Introduction, Lopez-Poveda et al. (2016b, 2017) demonstrated that BiCi users can show better speech-in-noise recognition with the MOC than with the STD strategy. For BiCI users, the benefits occur for spatially separated target and interferer sound sources. For unilateral CI users, the benefits occur when the implanted ear has the better acoustic SNR (**Fig. 2.7**). The MOC strategy, however, also had drawbacks: (1) it reduced the speech information in the ear with the worse acoustic SNR, which can potentially hinder intelligibility in unilateral listening when the implant ear had the worse acoustic SNR; and (2) the mutual inhibition between the pair of MOC processors decreased the overall stimulation levels and thus audibility, which can hinder intelligibility in bilateral or unilateral listening when the two CIs (or processors) have identical input signals.

Lopez-Poveda and Eustaquio-Martín (2018) reasoned that the tests of Lopez-Poveda et al. (2016b, 2017) were limited to an implementation of the MOC strategy with short (2 ms) time constants for the activation and deactivation of the inhibition. Such implementation and parameters disregarded aspects of the natural MOCR including the rather slow time courses for activation and deactivation of inhibition (Cooper and Guinan, 2003; Backus and Guinan, 2006), the possibility that the inhibition of BM responses be greater in apical than in basal cochlear regions (Lilaonitkul and Guinan, 2009; Aguilar et al., 2013), and the possibility that the largest MOCR inhibition occurs when the contralateral sound elicitor is one-half octave below the probe frequency (Lilaonitkul and Guinan, 2009). Furthermore, using a speech intelligibility model, Lopez-Poveda and Eustaquio-Martín (2018) predicted that the use of longer time constants for activation and deactivation of contralateral inhibition, combined with comparatively greater inhibition in the lower than in the higher frequency channels, can overcome the shortcomings of MOC processing and even improve the signal information in the ear with the worse acoustic SNR. In addition, they predicted no benefit of implementing a half-octave frequency offset in the contralateral control of inhibition.

---

[4] This chapter is based on the published paper: Lopez-Poveda EA, Eustaquio-Martín A, Fumero MJ, Gorospe JM, Polo R, Gutiérrez Revilla A, Schatzer R, Nopp P, Stohl JS. (**2020**). Speech-in-noise recognition with more realistic implementations of binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. *Ear and Hearing*. https://doi.org/10.1097/aud.0000000000000880

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 binaural advantage provided by MOC processing. Speech reception thresholds were measured 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. Speech reception thresholds 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 than in the higher frequency channels (MOC3) (see **Chapter 3** for details). Measurements with a slower MOC strategy with offset contralateral control of inhibition were not conducted because of time constraints and because, as explained above, 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), tests 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 CI users not previously tested on any of the strategies.

4.2. MATERIAL AND METHODS

4.2.1. Participants

Eight bilateral and two unilateral users of MED-EL CIs participated in the study (**Table 3.1**). 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). 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 below). One of the children (SA013) had been living in Scotland for the last four years but he spoke Spanish at home. All participants were reported to perform very well with their implants. Participant SA009 had not been wearing his left implant for a month just before the start of the study because the audio processor was damaged.

4.2.2. Stimuli

Speech reception thresholds were assessed as described in Section 3.4 using the Castilian Spanish version (Huarte, 2008) of the hearing-in-noise test (HINT) test (Nilsson et al., 1994) for a male target speaker. Speech reception thresholds were measured using fixed-level speech (at −20 dB FS) and varying the noise level adaptively. For reference, the speech level of −20 dB FS corresponds approximately to 70 dB SPL in MED-EL's clinical CI audio processors. 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, 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 ms before the sentence onset and ended 500 ms after the sentence offset and was gated with 50-ms cosine-squared onset and offset ramps.

4.2.3. Test conditions

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 3.1**), and in bilateral listening, involving listening with the two implants. Speech reception thresholds 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, $S_{-15}N_{15}$, $S_{-60}N_{60}$, $S_{-90}N_{90}$. 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.

4.2.4. Processing strategies

Stimuli were processed through the STD, MOC1, MOC2, and MOC3 sound-processing strategies prior to their presentation to participants. These strategies have been described in Section 3.2.

4.2.5. Order of testing

Unilateral listening tests were always administered first followed by bilateral listening tests. For each of the two listening conditions (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 standard deviation criterion (see Section 3.4), an additional SRT measurement was obtained after the full set of unilateral and bilateral tests was completed.

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 re-using the lists are discussed below.

4.2.6. Comparative analysis of STD and MOC output envelopes

In this section, the functioning of the tested strategies is described. The top part of **Figure 4.1** (panels **A** to **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 3.1**). For conciseness, output envelopes are shown only for channels #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 SSN and was located at −60° azimuth. The speech and noise had equal RMS levels at −20 dB FS (i.e., 0 dB SNR) and the noise started 500 ms before the speech onset, as in the SRT measurements. The bottom part of **Figure 4.1** (panels **M** to **Y**), shows the corresponding time course of the maplaw (or compression) *c* parameter (Eq. 3.1).

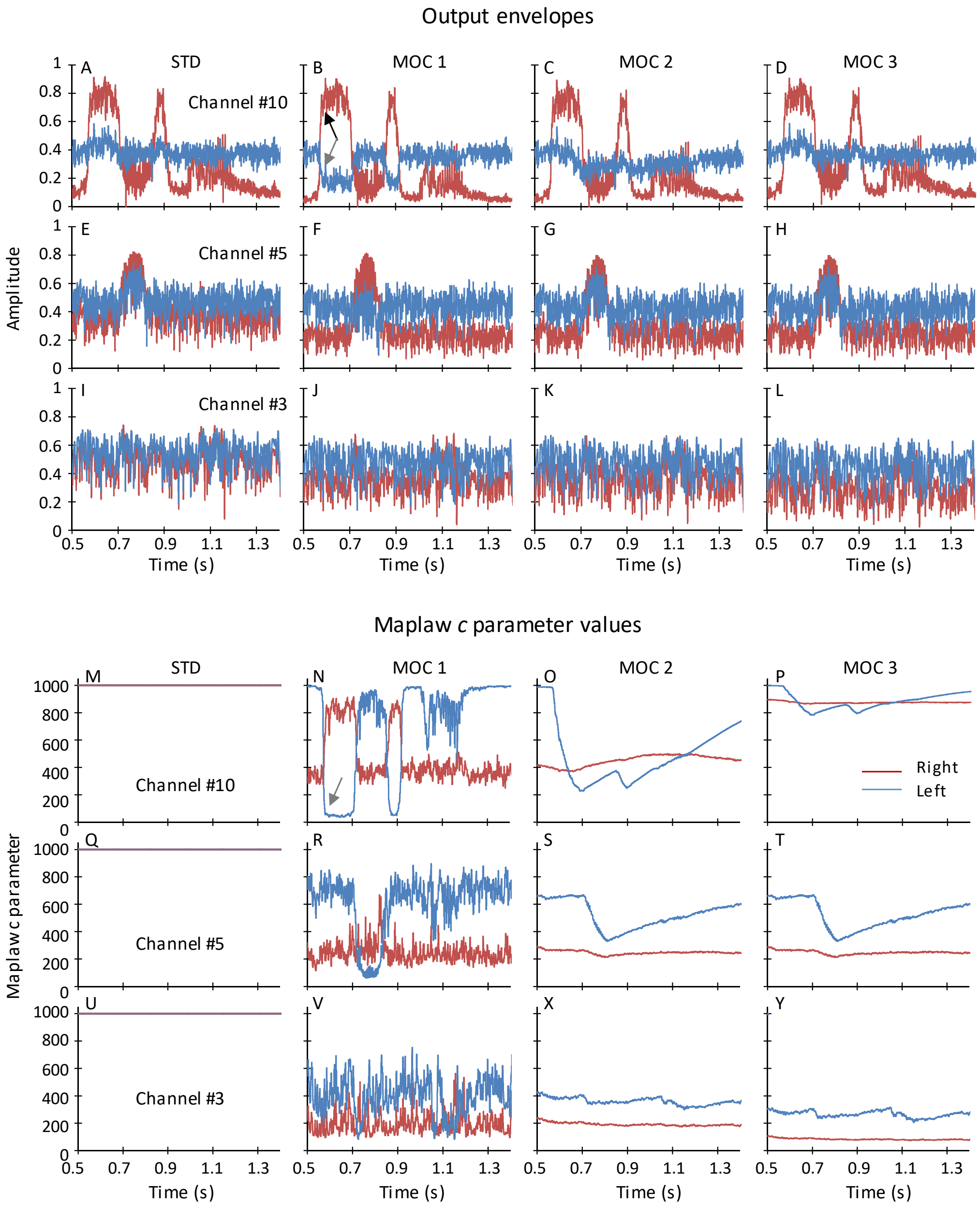

Figure 4.1. Example compressed envelopes (**top panels**) and maplaw values (**bottom panels**) for STD, MOC1, MOC2 and MOC3 strategies with 10 frequency channels. Data are shown only for three channels: channel #3 (bottom row), channel #5 (middle row), and channel #10 (top row) with center frequencies 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 ms 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). See the main text for details.

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 output envelope amplitude in a given frequency channel was larger in one ear (black arrow in **Fig. 4.1B**), the maplaw *c* parameter and thus the output amplitude, decreased in the corresponding contralateral frequency channel relative to the STD strategy (grey arrows in **Fig. 4.1B** and **Fig. 4.1N**). 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 (see Section 3.2.2).
4. For higher frequency channels (channel #10), which had larger bandwidths and thus produced higher output levels for broadband stimuli, inhibition was greater for MOC1 or MOC2 processors than for the MOC3 processors (i.e., the maplaw parameter was overall smaller in **Fig. 4.1N** or **Fig. 4.1O** than in **Fig. 4.1P**). 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. 3.2).
5. For lower frequency channels (channel #3) inhibition was greater for MOC3 than for MOC2 processors (i.e., the maplaw parameter was slightly smaller in **Fig. 4.1Y** than in **Fig. 4.1X**) because of bandwidth normalization.
6. For the normalization frequency channel (channels #5 in this example), the MOC2 and MOC3 processors had identical output envelopes (i.e., **Fig. 4.1G** was identical to **Fig. 4.1H**) and maplaw values (i.e., **Fig. 4.1S** was identical to **Fig. 4.1T**).

MOC processing can have several potential benefits over STD processing. To better understand them, **Figure 4.2** zooms in the output envelopes for channel #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 contralateral inhibition of the lower levels more than the higher levels (Lopez-Poveda et al., 2016). 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). Importantly, the reduction in vowel peaks was minimal in the ear closer to 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. 4.3**). 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. 4.1**). This is particularly true for the lower frequency channels, where speech envelope cues are more salient. As shown by the inset in **Figure 4.2A**, this can enhance the representation of the vowel envelope, which is the acoustic cue that most current cochlear implant 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 ILDs dynamically and on a channel-by-channel basis, as revealed by the fact that the maplaw values in **Figure 4.1** were different for the two ears.

Figure 4.2 also serves to illustrate some of the main differences across MOC processors. Compared to 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 4.2B**, 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. 4.3**). Additionally, 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.

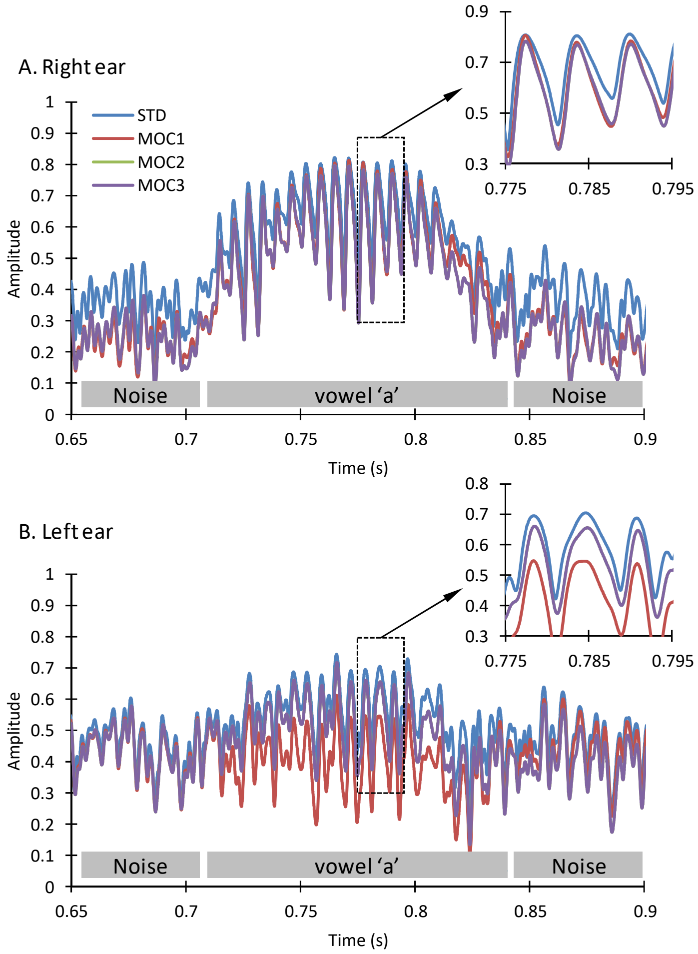

Figure 4.2. Zoomed-in view of the compressed envelopes for channel #5 shown in **Figure 4.1**. Each panel shows envelopes for the STD, MOC1, MOC2 and MOC3 strategies. Envelopes were identical for the MOC2 and MOC3, hence the overlap between corresponding traces. The gray rectangles near the abscissae 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.

Figure 4.3 summarizes of 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. Note 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 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 #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.

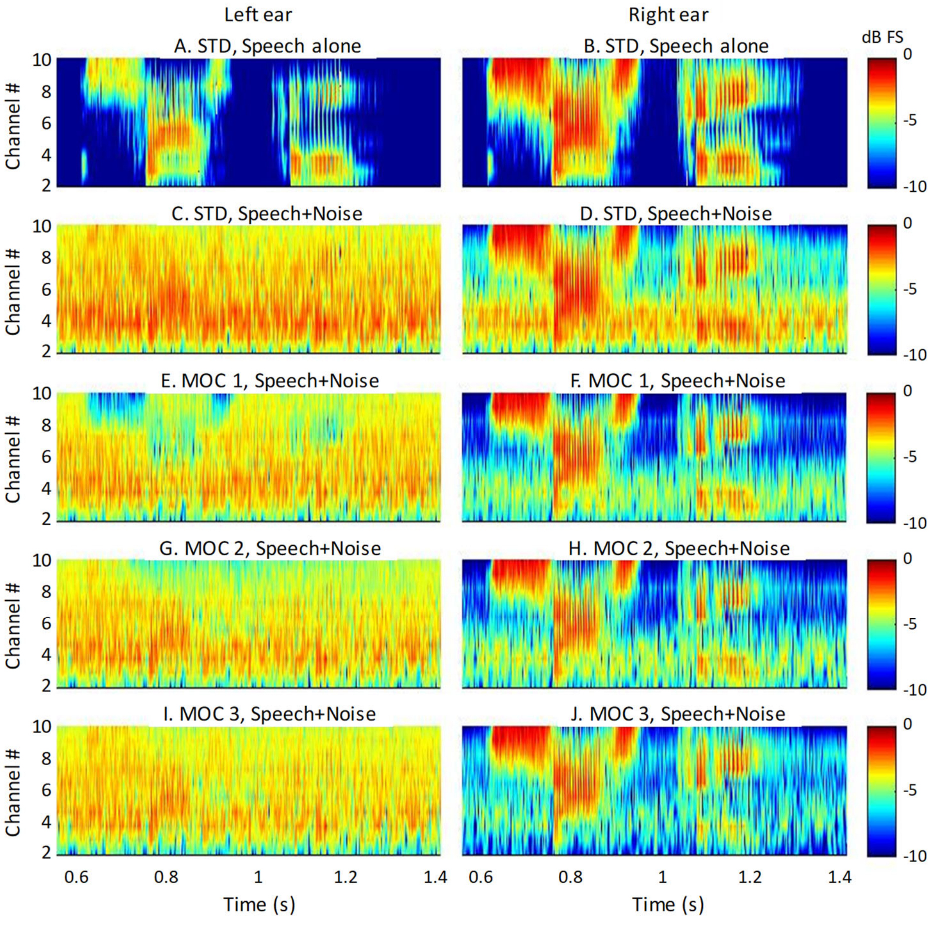

Figure 4.3. Output envelopes for STD, MOC1, MOC2 and MOC3 processors with 10 frequency channels. The stimulus was as in **Figure 4.1**. Each panel shows envelopes at the output of the back-end compression (or acoustic-electric maplaw) as a function of frequency channel number and time. Color illustrates 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).

MOC processing can have one additional benefit (relative to STD processing) not seen in the output envelopes (not seen in **Fig. 4.1**, **Fig. 4.2** or **Fig. 4.3**): 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 and Guinan, 1998; Lopez-Poveda, 2018).

4.3. RESULTS

In this section, the SRTs for the various MOC strategies are first compared with those for the STD strategy in unilateral and bilateral listening. Then, the potential advantage of using two ears versus one ear with the tested processing strategies is analyzed.

4.3.1. Speech reception thresholds in unilateral listening

The top row in **Figure 4.4** 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 4.4** Illustrate the SRT improvement or *benefit* (in decibels) relatively to the STD strategy provided by the MOC1, MOC2 and MOC3 strategies, respectively. The benefit was calculated as follows:

$$\mathrm{SRT_{benefit}}\ [\mathrm{dB}] = \mathrm{SRT_{STD}}\ [\mathrm{dB\ SNR}] - \mathrm{SRT_{MOC}}\ [\mathrm{dB\ SNR}] \quad , \qquad (4.1)$$

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 4.5** Shows group mean results.

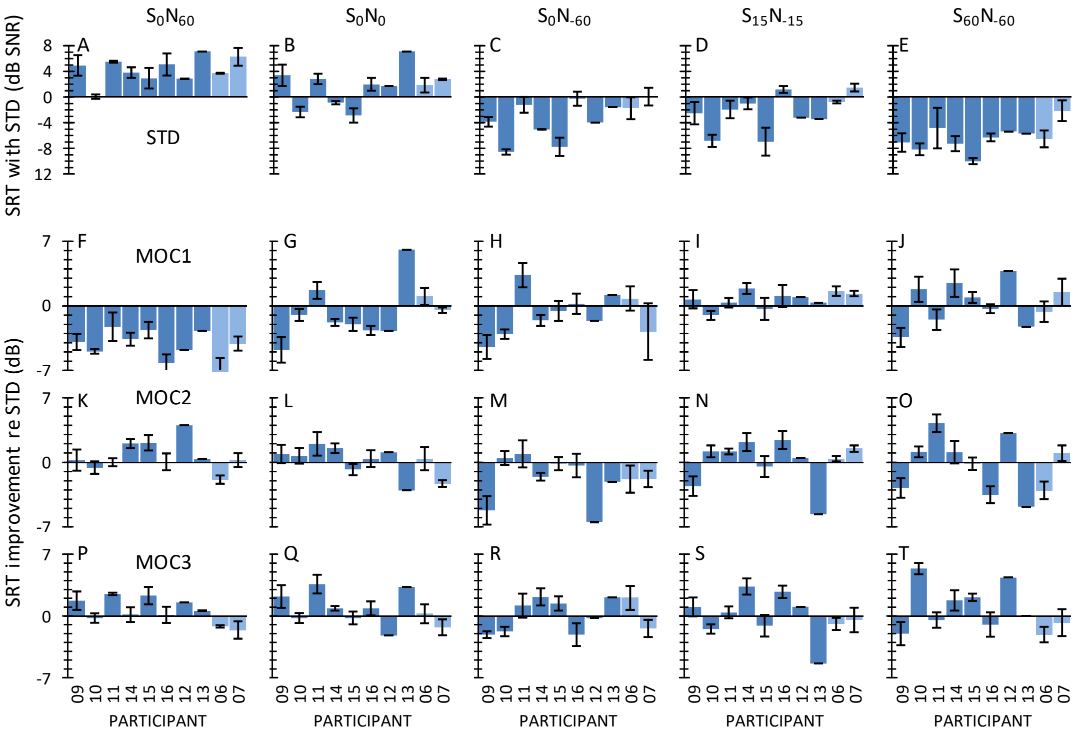

Figure 4.4. Intelligibility in unilateral listening for individual participants. **Row 1 (panels A to E)**. Speech reception thresholds for the STD strategy. Each panel is for a different spatial configuration of the speech and noise sources, as indicated at the top. **Rows 2 to 4 (panels F to T)**. Speech-reception-threshold 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 one standard error of the mean.

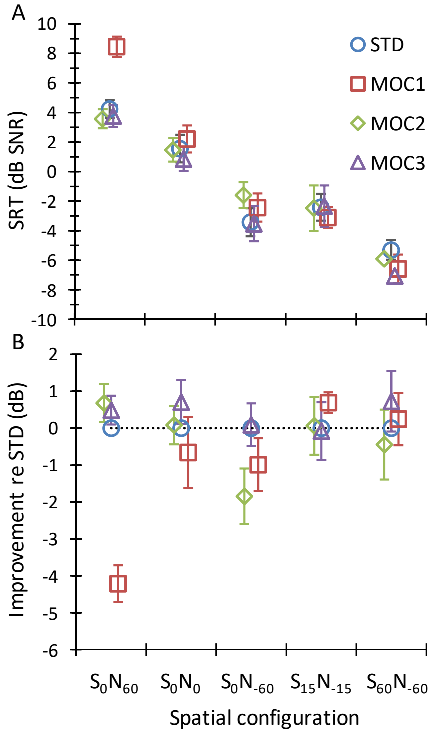

Figure 4.5. Group-mean intelligibility scores in unilateral listening. **A.** Mean SRTs for each strategy (as indicated by the inset) and spatial configuration (as indicated in the abscissa). Each point is the mean for eight bilateral and two unilateral CI users. **B.** Mean SRT improvement for the MOC strategies relative to the STD strategy. Error bars illustrate one standard error of the mean.

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 the participants (**Fig. 4.4F**). This is consistent with STOI simulations (see Fig. 5C in Lopez-Poveda and 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. 4.3C** and **Fig. 4.3E**). In contrast, SRTs were equal or better (up to 4 dB better for participant SA012) with the MOC2 than with the STD strategy (**Fig. 4.4K**), 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. 4.4P**). 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 about 4.2 dB worse with the MOC1 then with the STD strategy but slightly better (< 1 dB) with the MOC2 or MOC3 than with the STD strategy (**Fig. 4.5B**).

For speech and noise sources co-located 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. 4.4G**). This was expected based on earlier studies (Lopez-Poveda et al., 2016a) and STOI simulations (Fig. 5D in Lopez-Poveda and 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. 4.4Q**). On average, SRTs were slightly worse with the MOC1 than with the STD strategy but slightly better with the MOC2 or MOC3 than with the STD strategy (**Fig. 4.5B**).

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. 4.4H** and **4.4M**). However, some participants benefited from the MOC3 strategy (**Fig. 4.4R**). 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 and Eustaquio-Martín, 2018). The reason for the discrepancy between the present experimental result and the STOI prediction is uncertain. STOI disregards the 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 other did not. Altogether, there was no clear benefit or disadvantage of MOC processing compared to STD processing (see also the mean SRT improvement in **Fig. 4.5B**).

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 vs. –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} vs. $S_{15}N_{-15}$ (mean SRTs across participants and processors were 5.0, 1.5, –2.7, –2.5, –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 vs. MOC1) < 0.001; p(STD vs. MOC2) = 1.00; p(STD vs. MOC3) = 1.00; p(MOC1 vs. MOC2) < 0.001; p(MOC1 vs. MOC3) < 0.001; p(MOC2 vs. 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. 4.5**). For the other spatial configurations tested, the effect of strategy on SRT was not significant.

4.3.2. Speech reception thresholds in bilateral listening

Figure 4.6 shows individual results in bilateral listening. The layout is the same as **Figure 4.4**. 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 4.7** shows corresponding group mean results.

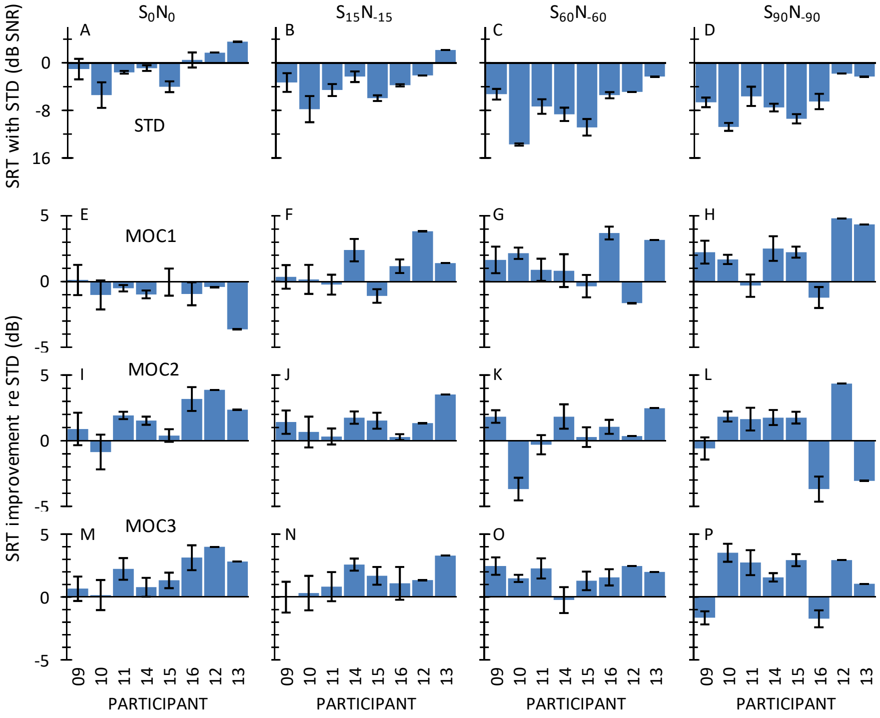

Figure 4.6. Intelligibility in bilateral listening for individual participants. The layout is the same as **Figure 4.4.**

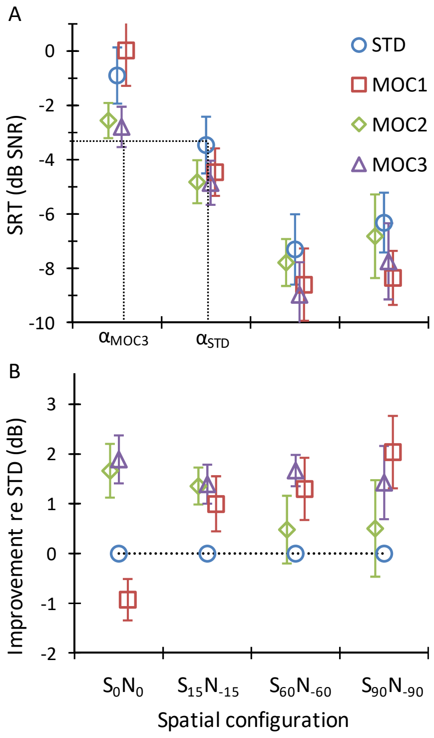

Figure 4.7. Group-mean intelligibility scores in bilateral listening. Each point is the mean for eight bilateral CI users. The layout is the same as Figure 4.5. The dashed lines in panel A illustrate that at a fixed SNR of about –3 dB, the angular separation between the speech and noise source (α) to achieve 50% correct sentence recognition would be narrower for the MOC3 than for the STD strategy ($\alpha_{MOC3} < \alpha_{STD}$); see main text for the details.

For co-located speech and noise sources (S_0N_0 condition), the MOC1 strategy was disadvantageous (the mean benefit was negative and equal to –0.9 dB, **Fig. 4.7**) 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. The exceptions were 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 [p(MOC1 vs. MOC2) = 1.00; p(MOC1 vs. MOC3) = 0.29; p(MOC2 vs. MOC3) = 0.50]. In addition, it revealed that the SRTs for the MOC2 and STD strategies were not significantly different from each other [p(STD vs. 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 vs. –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 vs. –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 (**Figure 4.7**).

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 $S_{60}N_{-60}$ vs. $S_{90}N_{-90}$ ($p=0.10$). 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 $S_{90}N_{-90}$.

4.3.3. Binaural advantage

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

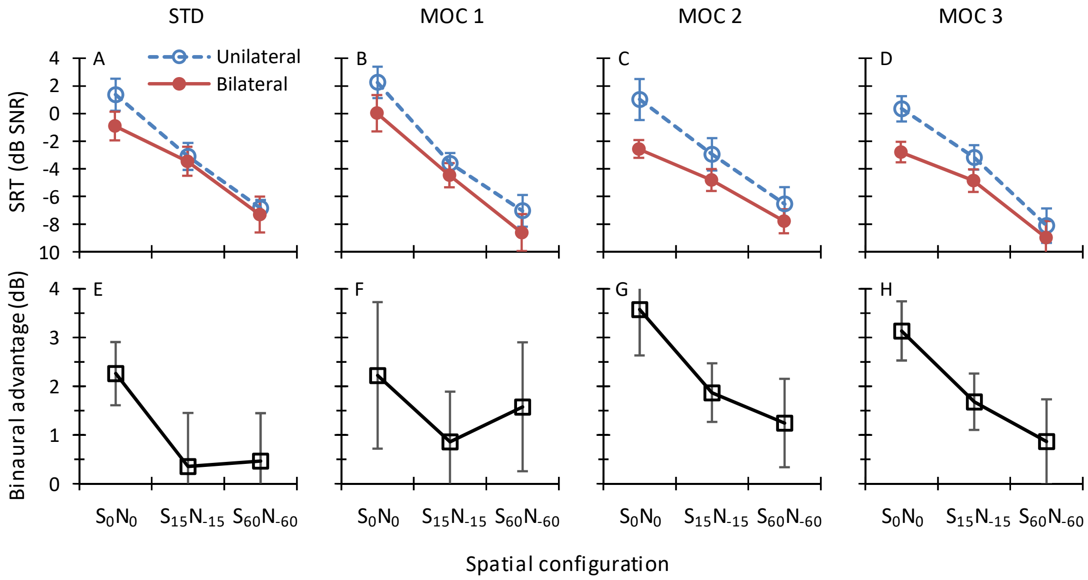

Figure 4.8. Top. Group mean SRTs in unilateral and bilateral listening. Each panel is for a different strategy, as indicated at the top of the panel. **Bottom**. Mean binaural advantage calculated as the difference in mean SRT for unilateral listening minus bilateral listening. Positive values indicate better (lower) SRTs when listening with two rather one ear. Error bars illustrate one standard error of the mean.

The top panels in **Figure 4.8** 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 4.8** 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 co-located speech and noise sources (S_0N_0 condition), bilateral listening tended to be more 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.

A RMANOVA was conducted to test for the effects of listening modality (unilateral *vs*. 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 4.1** shows the results. Significant effects are highlighted using bold font. Speech reception thresholds 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. Speech reception thresholds 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 co-located (S_0N_0, p = 0.007) or separated by 30 degrees ($S_{15}N_{-15}$: p = 0.017), but not when they were separated by 120 degrees ($S_{60}N_{-60}$: p = 0.21).

Table 4.1. 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. A separate test was conducted for each processing strategy (STD, MOC1, MOC2 and MOC3). Statistically significant effects are indicated using bold font.

Strategy	N	Listening condition	Spatial configuration	Interaction
STD	8	F(1,7) = 2.78, p = 0.139	**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	**F(2,14) = 106.22, p<0.001**	F(2,14) = 0.36, p = 0.700
MOC2	8	**F(1,7) = 10.36, p = 0.014**	**F(2,14) = 97.28, p<0.001**	**F(2,14) = 4.32, p = 0.034**
MOC3	8	**F(1,7) = 20.22, p = 0.003**	**F(2,14) = 88.06, p<0.001**	F(2,14) = 2.86, p = 0.091

Altogether, the present analysis demonstrates that only the MOC2 and MOC3 strategies produced a statistically significant binaural advantage; i.e., 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 4.8,** 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 the MOC3 (p=0.001) strategies, but not for the STD (p=0.061) or the MOC1 (p=0.336) strategies. 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.

4.4. DISCUSSION

Previous studies have shown that, compared to 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 2016a, 2017). The MOC1 strategy, however, produces equal or worse SRTs for co-located 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 than in the higher frequency channels (MOC3 strategy).

The main findings were:

(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. 4.7**). 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. 4.5**). This may seem inconsistent with a previous study that reported the MOC1 to be advantageous over the STD strategy in similar conditions (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 SSN 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. 4.5**). 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 co-located speech and noise sources, where the mean SRT was slightly worse (0.9 dB higher) for the MOC1 than for the STD strategy (**Fig. 4.7**). The MOC3 strategy, however, was advantageous over the STD strategy for all spatial configurations tested, including the co-located 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 co-located speech and noise sources) can be overcome by using slower control of compression combined with greater effects in the lower 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. 4.5**) and bilateral listening (**Fig. 4.7**) 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. 4.8**).

(7) The binaural advantage with the MOC2 and MOC3 strategies was significant for co-located (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. 4.8**).

Compared to 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.

4.4.1. 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 co-located speech and noise sources (S_0N_0) minus the SRT for spatially separated sources (see, for example, Fig. 4 in the review of Litovsky and Gordon, 2016). According to this definition, the data in **Figure 4.7** show that the mean spatial release from masking in bilateral listening for the $S_{60}N_{-60}$ vs. 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 co-located 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 the mean SRTs for the reference, co-located 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 correctly recognize 50% of the sentences with a smaller angular separation when using the MOC3 than when using the STD strategy. For example, the dashed lines in **Figure 4.7A** 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, it would be expected 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.

4.4.2. 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., Buss et al., 2008; Litovsky et al., 2006; Loizou et al., 2009; Schleich et al., 2004; Tyler et al., 2002) 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 summation 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. 4.8E**). In addition, for the S_0N_{-90} condition, Schleich et al. (2004) 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. (2004) supports the present findings and allow us to be optimistic that similar findings might be obtained in an eventual testing of the MOC strategies in the free field.

Compared to the STD strategy, the best MOC strategy (MOC3), and in general all MOC strategies, produced overall larger benefits in bilateral (**Fig. 4.7**) than in unilateral (**Fig. 4.5**) 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 #3) and the first block of bilateral listening tests (block #4), which were conducted consecutively. The pattern of results was also similar for the two last blocks of unilateral and bilateral listening tests (block #3 and block #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 4.3** and by Lopez-Poveda and Eustaquio-Martin (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. 4.5**), 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.

4.4.3. Limitations

Given the limited number of sentence lists in the HINT corpus, the sentence lists had to be used 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’ towards the end. As a result, the reported SRTs are probably lower than they would have been if the speech material had not been used repeatedly. We are confident, however, that re-using the sentences did not contribute to the reported differences in SRTs across strategies (or spatial configurations) because anyone 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 and Shannon, 1998; Theelen-van den Hoek et al., 2016). Here, compression in the STD processor (i.e., the value of parameter *c* in Eq. 3.1) was set to a (fixed) value that was not always the value used by the participants in their clinical processors (see **Chapter 3**). 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., 2016) and single-talker maskers (Lopez-Poveda et al., 2017), even when compression in 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-Martin, 2018). Third, **Figure 4.9** shows that STOI scores (computed as described by Lopez-Poveda and Eustaquio-Martin, 2018) are equal or higher for the MOC3 strategy than for a 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.

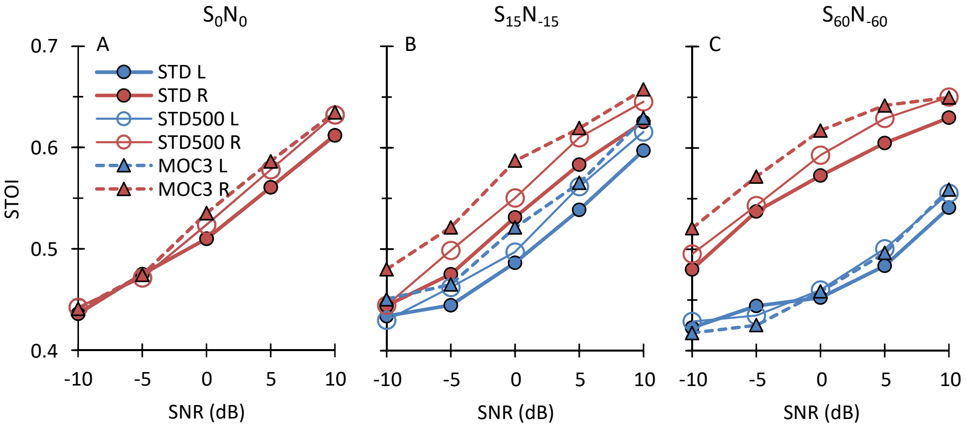

Figure 4.9. Comparison of STOI scores for the present STD and MOC3 strategies against scores for a STD strategy with *c*=500, the value typically used by the participants in their clinical audio processors. Each panel shows scores for the left (L) and right (R) ears (blue and red traces, respectively) 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.

4.5. CONCLUSIONS

The SNR at 50% 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 STD), or the use of binaurally coupled processors with contralaterally controlled dynamic compression inspired by the medial olivocochlear reflex (the MOC strategy). Three versions of the MOC strategy were tested: the MOC1, MOC2 and MOC3 strategy. The main conclusions are:

(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 co-located sources. The MOC3 strategy, however, was advantageous over the STD strategy for all spatial configurations tested, including the co-located 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 co-located 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.

5.

LATERALIZATION OF VIRTUAL SOUND SOURCES WITH A BINAURAL COCHLEAR IMPLANT SOUND-CODING STRATEGY INSPIRED BY THE MOCR[5]

5.1. INTRODUCTION

As reviewed in the General Introduction, inadequate coding of binaural cues by the CI audio processors likely contributes to the poorer localization performance of BiCI users compared to normal-hearing listeners (Dorman et al., 2016). One specific factor that can potentially degrade the coding of ILD cues, needed to localize sound sources, is the use of independent compression in the two audio processors of a BiCI user. As illustrated in schematic form in **Figure 5.1A**, the application of independent AGC and/or acoustic-to-electric maps to the two ears can compress (reduce) the head-shadow ILDs and thus hinder the localization of sound sources in the horizontal plane (e.g., Dorman et al., 2014; Ricketts et al., 2006; Wiggins and Seeber, 2011). Indeed, BiCI users can localize sounds more accurately with binaurally linked rather than with independent AGC in their two devices (Potts et al., 2019).

The MOC strategy, by using binaurally coupled back-end compression, can also enhance the head-shadow ILDs in each frequency channel of processing (see Fig. 2 in Lopez-Poveda et al., 2016a), thus the overall ILDs, relative to that available with two independently functioning processors (the STD approach). **Figure 5.1B** illustrates the mechanism in schematic form. Insofar as BiCI users rely mostly on ILD cues for localization, it seems possible that sound source localization in the horizontal plane may be better with the MOC than with the STD strategy. The main aim of the present study was to investigate this possibility using virtual acoustic stimuli. On the other hand, the ILDs delivered by the MOC strategy depend on the amount of contralateral inhibition of compression, which can be set using parameters. A second aim was to compare sound lateralization performance with various implementations of the MOC strategy designed to reflect more or less realistically the inhibitory characteristics of the natural contralateral MOCR (Lopez-Poveda and Eustaquio-Martín, 2018).

---

[5] This chapter is based on the published paper: Lopez-Poveda EA, Eustaquio-Martín A, Fumero MJ, Stohl JS, Schatzer R, Nopp P, Wolford RD, Gorospe JM, Polo R, Gutiérrez Revilla MA, Wilson BS. (**2019**). Lateralization of virtual sound sources with a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. *Hearing Research* **379**:103-116. https://doi.org/10.1016/j.heares.2019.05.004

To address these aims, twelve BiCI users were asked to localize noise tokens in a virtual horizontal plane with the STD strategy and three different implementations of the MOC strategy (MOC1, MOC2 and MOC3).

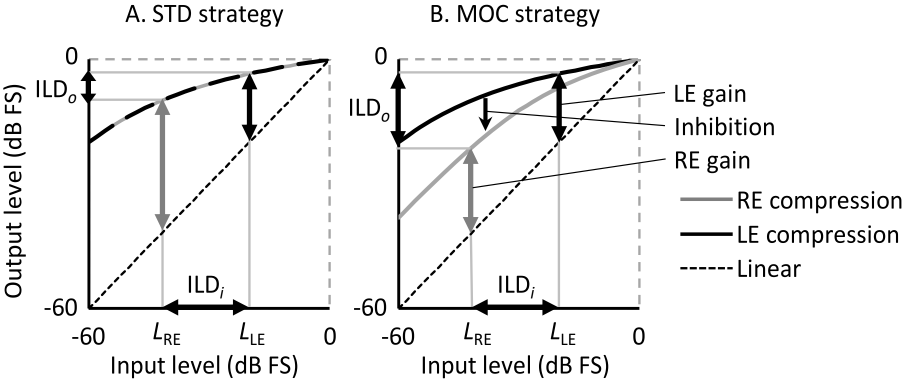

Figure 5.1. Schematic interaural level difference at the output (ILD$_o$) for a given interaural level difference at the input (ILD$_i$), with two different back-end compression schemes. **A.** with equal and independently functioning compressors at the two ears (STD strategy). **B.** With binaurally coupled compressors as in the MOC strategy. Each panel illustrates the ILD$_o$ (double headed arrow on the ordinate) for a hypothetical sound source located in the free field on the left side of the head; i.e., when the stimulus (input) levels is greater on the left than on the right ear ($L_{LE}>L_{RE}$). With the STD strategy (**A**), the compression function would be identical for the right and the left ear, hence the overlap between the black and gray curves. Because the input stimulus level is smaller on the right ear, the right-ear compressor applies more gain (re linear) to the stimulus than the left-ear compressor does. As a result, ILD$_o$<ILD$_i$. With the MOC strategy (**B**), the output level is larger for the left than for the right ear processor. Therefore, the left ear output inhibits the right-ear compressor more than other way around. This would turn the right-ear compressor more 'linear' with minimal or no change in the left-ear compressor. As a result, ILD$_o$ would be larger with the MOC than with the STD strategy. Note that in this example, the compression functions were calculated using Eq. 3.1, and that the input and output levels are in logarithmic scales in dB FS, where 0 dB FS corresponds to a peak amplitude at 1, which itself corresponds to an electrical current at maximum comfortable loudness (MCL). RE: right ear; LE: left ear.

5.2. METHODS

5.2.1. Participants

Twelve users of bilateral MED-EL CIs participated in the study (**Table 3.1**). Three of them (ME115, ME131 and ME132) were tested at the MED-EL US Laboratory (North Carolina, USA), and nine of them were tested at the University of Salamanca (SA004, SA005, SA008, SA009, SA010, SA011, SA014, SA015 and SA016).

5.2.2. Stimuli

Localization was assessed as described in Section 3.5. Stimuli consisted of digitally generated Gaussian noise bursts bandpass filtered between 125 and 6000 Hz with a fourth-order (North Carolina) or first-order (Salamanca) Butterworth filter to achieve the desired bandwidth. The noise bursts had a duration of 200 ms and were gated with 20-ms (North Carolina) or 50-ms (Salamanca) raised-cosine onset and offset ramps. A linear gain was applied to the noise bursts to achieve the desired presentation level of −20 dB FS (dB relative to a peak amplitude at unity). For reference, this level corresponds approximately to 70 dB SPL in MED-EL's clinical CI audio processors. For the North Carolina participants, the stimulus level was randomly roved by up to ±2 dB across stimulus presentations; for the Salamanca participants, the stimulus level remained constant across stimulus presentations. The potential implications of this approach are discussed later.

The level-adjusted noise bursts were preceded and followed by silence periods with a duration of 20 ms, making the stimulus duration equal to 240 ms. Stimuli were presented for sound sources at 0° elevation and for 11 azimuthal angles from –75° to 75° separated by 15°. This stimulus choice was intended to facilitate a comparison between the present results (which were for a virtual acoustic setting using non-individualized HRTFs and with experimental processing strategies) with previous reports of the performance of BiCI users tested in the free-field with their own clinical devices (Dorman et al., 2014; 2016; see the Discussion section).

5.2.3. *Processing strategies*

The level-adjusted, HRTF-filtered noise bursts were processed through the STD, MOC1, MOC2 and MOC3 strategies before they were presented to the BiCI participants via direct stimulation. The processing strategies have been described in detail in Section 3.2.

5.2.4. *Data analyses*

Response matrices were generated by plotting the reported against the actual azimuthal angles. Localization accuracy was quantified using the RMS angle of error (E_{RMS}), calculated as:

$$E_{RMS} = \sqrt{\sum_{i=1}^{N} \frac{(X_i - Y_i)^2}{N}} \quad , \qquad (5.1)$$

where X_i and Y_i denote the actual and reported azimuthal angles for the i-th stimulus presentation, and N is the total number of presentations ($N = 88$). Localization accuracy was also quantified using the Pearson correlation coefficient between actual and reported azimuthal angles, R_{XY}. These two-performance metrics are complementary. The correlation coefficient can be advantageous over the RMS angle error when the reported location is systematically to the left or the right of the actual location due to potentially inadequate binaural loudness balance (e.g., see Fig. 3 in Tyler et al., 2006). Conversely, the correlation coefficient is insensitive to potential systematic lateralization bias (i.e., to vertical offsets in the response matrices) that might increase $\varepsilon_{\mathrm{RMS}}$. Both R_{XY} and $\varepsilon_{\mathrm{RMS}}$ are commonly used to quantify accuracy in localization studies (e.g., Majdak et al., 2013; Marmel et al., 2018).

5.3. RESULTS

In this section, the level cues provided by the different processing strategies and their time course are first analyzed. Then, the localization scores for the (originally proposed) MOC1 strategy with those for the STD strategy (aim 1 of the study) are compared. Lastly, localization scores for the various implementations of the MOC strategy (MOC1, MOC2 and MOC3) with those for the STD strategy are compared.

5.3.1. Level cues provided by the STD and MOC strategies

Figure 5.2 shows output envelopes for a sound source located at –60° azimuth from STD, MOC1, MOC2 and MOC3 processors with 10 frequency channels, the typical number of channels for the present participants (**Table 3.1**). For conciseness, output envelopes are shown for three channels only: channel #3 (bottom row), #5 (middle row) and #10 (top row), with center frequencies of 501, 1159 and 7230 Hz, respectively. The left and middle columns illustrate output amplitudes for the left ear and right ear, respectively, and the right-most column illustrates the difference in output amplitude between the left and the right ears (note that this is different from the output ILD, which is discussed below). To better illustrate the effect of contralateral inhibition in the MOC strategies, the stimulus consisted of ten pure tones equal in amplitude and whose frequencies were approximately at the center of the processors' frequency channels. The stimulus was long enough (its duration was 2 s with 50-ms cosine-squared onset and offset ramps) to reveal the full inhibitory effects of the slower MOC strategies, and its overall level was set at –20 dB FS, thus equal to the level of the noise bursts used in the localization experiments. To facilitate the visualization of the different traces using different line styles, the output signals were first smoothed (using Matlab's smooth function) and then downsampled from 20 kHz to 40 Hz.

The figure illustrates the following:

1. For any given processing strategy and frequency channel, the output amplitude was greater for the left ear (left column in **Fig. 5.2**) than for the right ear (middle column in the **Fig. 5.2**). This is because the sound source was located on the left side of the head (at –60° azimuth) and the HRTF introduced a head-shadow ILD. The interaural amplitude difference (right panels in **Fig. 5.2**) was larger for channel #10 than for channels #3 or #5 because channel #10 was higher in frequency and the head-shadow ILD is greater at higher than at lower frequencies (e.g., Blauert, 1997; Lopez-Poveda, 1996).
2. In the right ear (the shadowed ear in this example; middle column in **Fig. 5.2**), the amplitude was always greater or equal for the STD strategy than for any of the MOC strategies. This is because in the MOC strategies, the ear with the largest output amplitude inhibits the ear with the smallest output amplitude more than the other way around. Because the output amplitude in this example was greater for the left ear than for the right ear, the left ear inhibited the right ear more than the other way around, which reduced the output amplitude more in the right than in the left ear.
3. Contralateral inhibition was faster for the MOC1 than for the MOC2 or MOC3 strategies. For the MOC2 or MOC3 strategies, it took approximately 1 s for the output amplitude in the right ear to achieve its asymptotic value. The different time course between the MOC strategies was related to using a faster time constant of contralateral inhibition in the MOC1 than in the MOC2 or MOC3 strategies.
4. The interaural amplitude difference (right column in **Fig. 5.2**) was equal or greater for any MOC strategy than for the STD strategy. This is because contralateral inhibition in the MOC strategies reduced the output amplitude in the right ear.
5. In the high frequency channel #10, the asymptotic interaural amplitude difference was greater with the MOC1 and MOC2 strategies than with the MOC3 strategy (**Fig. 5.2C**). The opposite was true for channel #3 (**Fig. 5.2I**). For channel #5 (**Fig. 5.2F**), all three strategies produced equal amount of contralateral inhibition. The different effect of MOC strategies

on the interaural amplitude difference was related to using (or not) bandwidth normalization (Eq. 3.2).

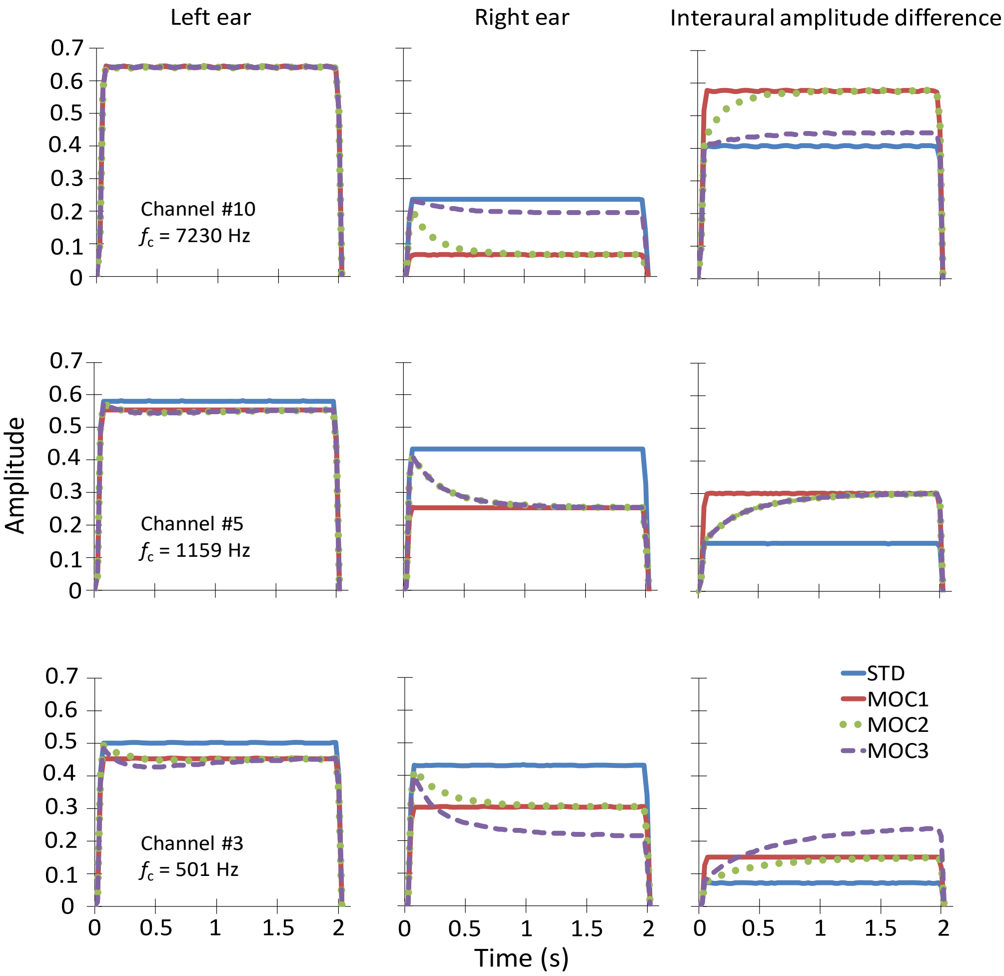

Figure 5.2. Example output signals (compressed envelopes) and interaural amplitude difference for STD, MOC1, MOC2 and MOC3 processors with 10 frequency channels, and for a sound source located at −60° azimuth. The MOC3 strategy was implemented with $BW_{\mathrm{ref}} = BW_{\#5}$. The stimulus was a ten-tone complex (2 s in duration with 50-ms onset and offset ramps), with all tones having identical input level. The overall stimulus level was −20 dB FS. Each row shows signals for a different frequency channel with center frequencies (f_c) of 501 Hz (channel #3, bottom row), 1159 Hz (channel #5, middle row), and 7230 Hz (channel #10, top row). **Left column.** Amplitude at the output of the left ear processor. **Middle column.** Amplitude at the output of the right ear processor. **Right column.** Difference in output amplitude between the left and the right ear. Each panel illustrates four traces (one per processing strategy), as indicated by inset. See main text for details.

Figure 5.3 illustrates the RMS output level (computed over the whole stimulus duration and expressed in dB FS) and the ILD (in dB) for each frequency channel and for various processors with 10 frequency channels. In this example, stimuli were identical (200-ms, wideband noise bursts) as those used in the experiments and the sound source was located at −60° azimuth. The top and middle panels illustrate output levels for the left and right ears, respectively; the bottom panel illustrates the ILD calculated as $20\times\log_{10}(O_{LE}/O_{RE})$, with O_{LE} and O_{RE} denoting the RMS output amplitudes (in linear units) at the left and right ears, respectively. To illustrate the effect of compression, **Fig. 5.3** also shows the output levels and ILDs for a 'linear' STD strategy without back-end compression [achieved by setting c = 1e−10 in Eq. (3.1)]. The peak and valleys of the STD-LIN trace reflect the stimulus spectrum combined with the spectral shape of the HRTF, the high-pass pre-emphasis filter, and the filter bank. The figure shows that for all strategies and channels, the output level was greater or equal for the left ear than for the right ear. This is because the sound source was located on the left side of the head and the HRTF introduced a head-shadow ILD. In addition, the output level was greater for any strategy than for STD-LIN because all four strategies (STD, MOC1, MOC2 and MOC3) applied back-end compression that amplified the lower input levels more than the high input levels. Compression, however, reduced the spectral contrast at each ear as well as the ILD (**Fig. 5.3C**). It was noted that of the four test strategies, the MOC1 strategy provided spectral contrast and ILDs that were most similar to the values that would be available without the 'detrimental' effects of compression (depicted as STD-LIN in **Fig. 5.3**). In addition, it was noted that the MOC2 and MOC3 strategies provided similar output levels, ILDs and spectral contrast as the STD strategy did because the stimulus was shorter (200 ms) than the time required for full activation of contralateral inhibition in the MOC2 and MOC3 strategies (**Fig. 5.2**).

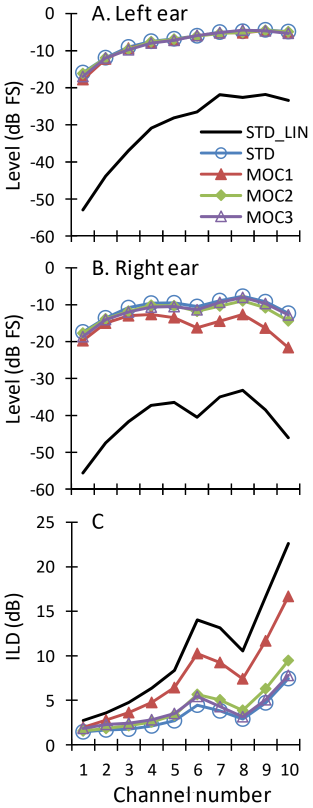

Figure 5.3. Output level (in dB FS) at the left ear (**A**), right ear (**B**), and ILD (**C**) as a function of channel number for STD, MOC1, MOC2, and MOC3 sound processors with 10 frequency channels. Also shown are the levels and ILD for a linear STD processor with minimal back-end compression (STD_LIN). The stimulus was a 200-ms wideband noise burst identical to those used for testing and the source was located at −60° azimuth. The MOC3 strategy was implemented with $BW_{\mathrm{ref}} = BW_{\#5}$.

Figure 5.4 illustrates the overall output levels at each ear as well as the ILD as a function of azimuth angle for STD, MOC1, MOC2, and MOC3 strategies with 10 frequency channels. Stimuli were identical (200-ms, wideband noise bursts) as those used in the experiments and were vocoded (using noise carriers) through the corresponding processing strategy. The vocoder has been described elsewhere (Lopez-Poveda and Eustaquio-Martín, 2018). For reference, the figure also shows the stimulus levels at the input of the processors (i.e., after HRTF filtering), depicted as HRTF. Because the MOC strategies are sensitive to stimulus levels (see Lopez-Poveda, 2015), results are shown for stimulus levels of −40, −30, −20 and −10 dB FS, as indicated at the top of each column. (Note that the present experiments with BiCI users were conducted with stimuli around −20 dB FS).

Figure 5.4 illustrates the following:

(1) For all strategies, the overall output levels at each ear increased gradually with increasing stimulus level (i.e., levels increased from the left-most to the right-most panels in the top and middle rows of **Fig. 5**.**4**). However, the difference between the input (HRTF) and the output levels decreased with increasing stimulus level (i.e., the length of the vertical arrows in the top and middle rows of **Fig. 5.4** decreased from left to right). This is because back-end compression (Eq. 3.1) amplified the lower input levels more than the higher input levels.

(2) At each ear, the MOC1 strategy produced the steepest level azimuth functions, the more similar in slope to the corresponding HRTF functions, and the more constant in slope across the range of stimulus levels tested (i.e., the dashed lines and the filled triangles in the top and middle rows had identical or very similar slopes from left to right). The STD, MOC2 and MOC3 strategies, by contrast, produced level-azimuth functions that became gradually shallower as the stimulus level increased (their slope decreased from left to right in **Fig. 5.4**). In other words, the MOC1 strategy preserved to a larger extent the monaural HRTF level localization cues across the range of stimulus levels tested. For the STD strategy, these monaural level cues decreased gradually with increasing sound level because compression enhanced the lower input levels in the shadowed ear more than the higher input levels in the ear closer to the sound source. While the contralateral inhibition of compression used in all MOC strategies can theoretically preserve those monaural cues, only the MOC1 strategy preserved those cues because only for this strategy was contralateral inhibition maximally active over the virtually full stimulus duration (i.e., as noted earlier and in **Fig. 5.2**, the stimulus duration was shorter than the time required for full activation of contralateral inhibition in the MOC2 or MOC3 strategies).

(3) The MOC1 strategy produced the largest ILD, the closest to the HRTF ILD, and the more constant across the range of stimulus levels tested. By contrast, the ILD was comparable for the MOC2, MOC3 and STD strategies, was smaller than the HRTF ILD, and nearly halved as the stimulus level increased from −40 to −10 dB FS. The ILD was largest for the MOC1 strategy because only for the MOC1 strategy was contralateral inhibition maximally active over virtually the full stimulus duration (see **Fig. 5.2**).

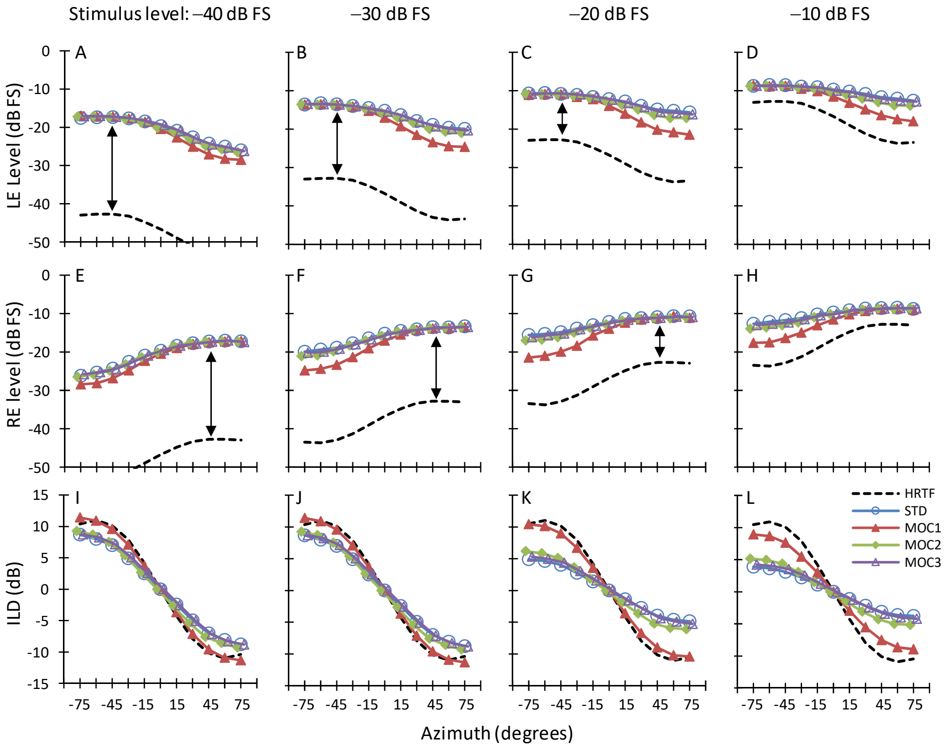

Figure 5.4. Overall output level at the left ear (LE, top row), right ear (RE, middle row), and ILD (bottom row) as a function of azimuth location for noise-vocoded stimuli with STD, MOC1, MOC2, and MOC3 sound processors with 10 frequency channels. The MOC3 strategy was implemented with BW_{ref} = $BW_{\#5}$. Also shown are the amplitudes at the input of the processors and the corresponding ILDs, depicted as HRTF. Each column shows results for a different stimulus level, from –40 to –10 dB FS, as indicated at the top of the column.

5.3.2. Localization with the MOC1 and STD strategies

Neither the angle error nor the correlation coefficient for each individual participant from the smaller North Carolina group (N = 3) were outside the mean plus or minus two standard deviations interval for the more numerous Salamanca group (N = 9). Furthermore, the North Carolina and Salamanca groups were not significantly different in mean angle error or correlation coefficient with either the STD or the MOC1 strategies (two-tailed Student t tests for unequal sample sizes with unequal variances produced *p* values of 0.45 for the difference in mean angle error between the two groups with the STD strategy, 0.13 for the difference in mean angle error with the MOC1 strategy; 0.56 for the difference in mean correlation coefficient with the STD strategy; and 0.95 for the difference in mean correlation coefficient with the MOC1 strategy). This justified analyzing the data for the Salamanca and North Carolina participants jointly.

Figure 5.5 shows example response matrices for two example participants: the ‘best’ overall performer with the smallest angle errors (SA004, top panels) and a typical performer with angle error scores close to the group mean scores (SA014, bottom panels).

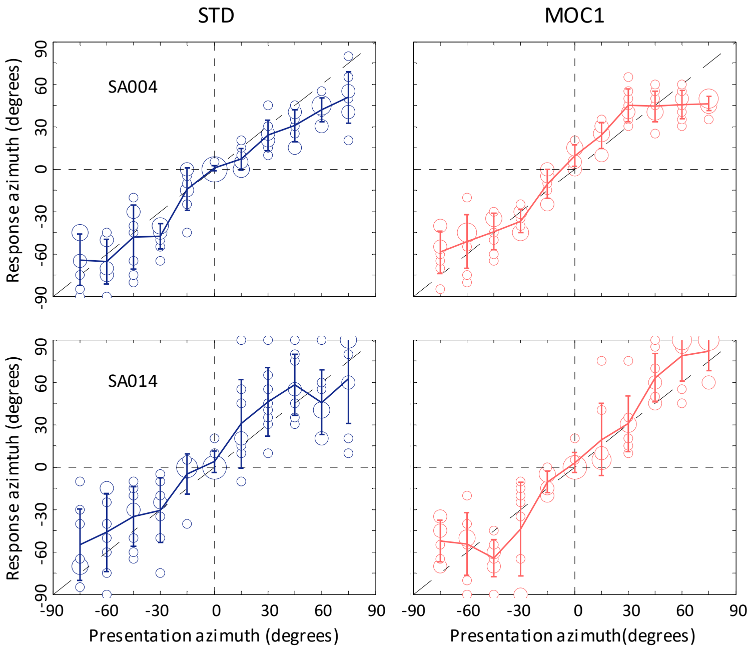

Figure 5.5. Example localization matrices for two BiCI users: SA004 (top) and SA014 (bottom). A pair of matrices is shown for each participant: one for the STD strategy (left) and one for the MOC1 strategy (right). Within each matrix, the reported azimuth is shown as a function of presented azimuth angle. Eight stimuli were presented for each azimuth angle. The size of each point is proportional to the number of responses at the corresponding angle. Lines show the mean reported angle for every actual angle and error bars illustrate one standard deviation.

Figure 5.6 illustrates individual and group mean localization angle error scores (ε_{RMS}, Eq. 5.1) for the MOC1 and STD strategies. Chance performance for the North Carolina and Salamanca setups (calculated by assessing random localization performance) was approximately 64° and 70°, respectively. All participants performed better than chance. For eight of the 12 participants, the angle error was smaller for the MOC1 than for the STD strategy. For participants SA005, SA009, SA011, and SA015, the angle error was comparable for the two strategies. Kolmogorov-Smirnov tests (with Lilliefors correction) revealed that angle error scores for the STD and MOC1 strategies each conformed to a normal distribution ($p>0.200$), thus it was justified to use parametric statistical tests to compare angle error score with the MOC1 and STD strategies. The group mean angle error score was smaller for the MOC1 (mean ± s.d. = 22.7°±3.6°) than for the STD strategy (25.3°±4.1°) and the difference was statistically significant (two-tailed, paired Student's *t*-test, $p = 0.0015$, $N = 12$).

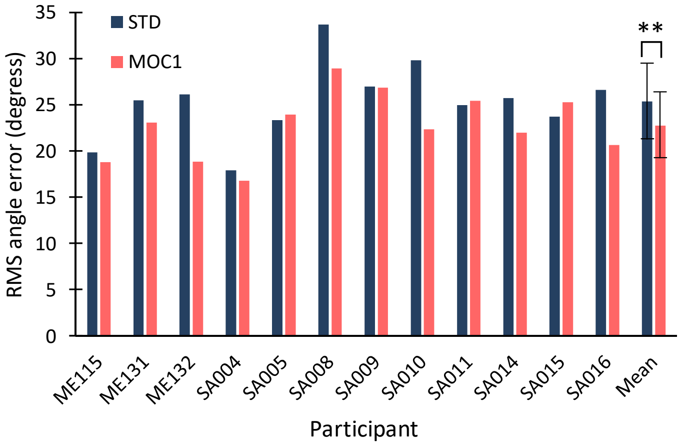

Figure. 5.6. Angle error for the MOC1 and STD strategies, as indicated by the inset. Results are shown for each individual participant and the mean across participants. Lower values indicate better performance. Error bars for the mean scores depict one standard deviation. **: *p* ≤ 0.01.

Figure 5.7 shows the correlation coefficient (R_{XY}) between the actual and reported azimuth for the MOC1 and the STD strategies for each individual participant and the mean across participants. Nine participants showed higher (better) correlation coefficients with the MOC1 than with the STD strategy and three participants (SA004, SA005 and SA008) showed similar correlation coefficients with the two strategies. Kolmogorov-Smirnov tests (with Lilliefors correction) revealed that the correlation coefficients for the STD and MOC1 strategies each conformed to a normal distribution (p>0.200). The group mean correlation coefficient was higher (better) with the MOC1 (mean ± s.d. = 0.92 ± 0.024) than with the STD strategy (0.89 ± 0.037) and the difference was statistically significant (two-tailed, paired Student's *t*-test, p = 0.005, N = 12).

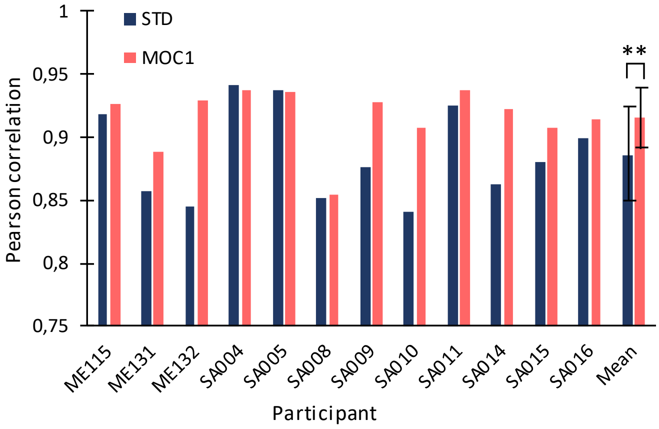

Figure 5.7. Correlation between presentation and response azimuth for the MOC1 and STD strategies, as indicated by the inset. Results are shown for each individual participant and the mean. Higher values indicate better performance. Note that the ordinate scale starts at 0.75 rather than zero to better show small differences. Error bars for the mean scores illustrate one standard deviation. **: *p* ≤ 0.01.

Figure 5.8A allows a comparison of group mean angle error scores for the MOC1 and the STD strategy for each azimuth location. The two strategies produced similar errors (within ±2°) for azimuths at or near ±30° (**Fig. 5.8B**). The MOC1 strategy, however, tended to improve lateralization for virtually every other azimuth, particularly for sources near the midline (i.e., for azimuths between –15°and +15°) and on the far sides (i.e., for azimuths ≥+60° or ≤-60°). A RMANOVA revealed a statistically significant effect of processing strategy [F(1,11)=10.52, p=0.008]. However, neither the effect of azimuth angle [F(10,110)=1.37, p=0.220] nor the interaction between processing strategy and azimuth angle were statistically significant [F(10,110) = 0.85, p = 0.581].

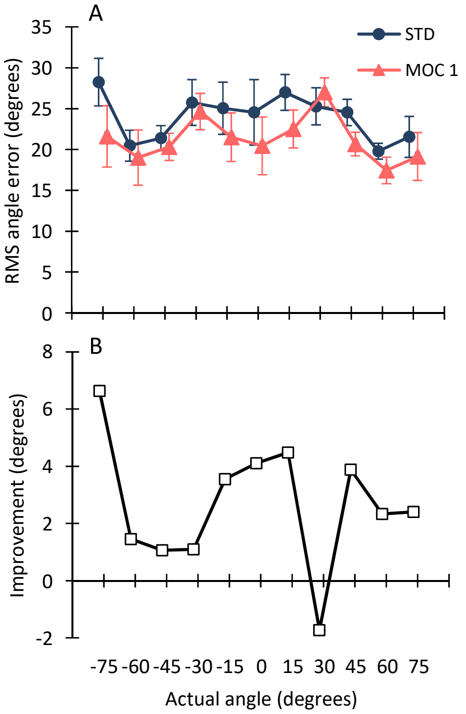

Figure 5.8. A. Mean RMS error angle for each azimuth location and for the STD and MOC1 strategies. Error bars illustrate one standard error of the mean (N = 12). Data points have been slightly displaced horizontally to reduce overlap. **B.** Mean localization improvement with the MOC1 strategy, calculated as the mean angle error for the STD strategy minus the mean angle error for the MOC1 strategy, all in degrees.

5.3.3. Localization with the STD, MOC1, MOC2, and MOC3 strategies

Seven of the 12 participants (SA008, SA009, SA010, SA011, SA014, SA015, and SA016) were tested with the STD, MOC1, MOC2 and MOC3 strategies. **Figure 5.9** allows a comparison of localization performance with the different strategies. The trends were different for different participants. The worst performance occurred for participant SA008 with the MOC3 strategy (angle error = 35°, correlation = 0.75). This was probably due to two factors. First, SA008 was the worst performer overall, regardless of the strategy. Second, in implementing the MOC3 strategy, BW_{ref} was made equal to $BW_{\#7}$ for participant SA008, while it was approximately equal to the BW of the median channel for all other participants (see Methods). Participant SA008 had 10 active electrodes and so normalizing to $BW_{\#7}$ probably caused excessive inhibition that compromised audibility, which may have degraded his performance. For these reasons, MOC3 scores for participant SA008 were omitted from the mean values in **Figure 5.9** and from the following statistical analyses.

Kolmogorov-Smirnov tests (with Lilliefors correction) showed that angle error scores conformed to a normal distribution for all four strategies (mean ± s.d. for STD: 27.4°±3.1°, *p*=0.150; MOC1: 24.5°±2.7°, *p*>0.200; MOC2: 27.2°±3.8°, *p*=0.135; MOC3: 24.6°±2.5°, *p*>0.200), thus it was justified to use a RMANOVA to test for the effect strategy on angle error score. The RMANOVA test revealed no significant effect of processing strategy on angle error [F(3,15)=1.49, *p*=0.26].

Kolmogorov-Smirnov tests (with Lilliefors correction) revealed that the correlation coefficient conformed to a normal distribution for the STD (0.88±0.026, *p*>0.200), MOC2 (0.88±0.036, *p*>0.200), and MOC3 (0.89±0.024, *p*>0.200) strategies, but not for the MOC1 strategy (0.91±0.025, *p*=0.011). A Friedman test revealed a statistically significant difference in correlation between actual and reported azimuth depending on the strategy [$\chi^2(3)=9.343$, *p*=0.025]. Post-hoc pairwise analysis with Wilcoxon signed-rank revealed a trend for better (higher) correlation with the MOC1 than with any other processing strategy (STD vs. MOC1: Z=–2.521, *p*=0.012; STD vs. MOC2: Z=–0.140, *p*=0.889; STD vs. MOC3: Z=–0.507, *p*=0.612; MOC1 vs. MOC2: Z=–2.240, *p*=0.025; MOC1 vs. MOC3: Z=–2.197, *p*=0.028; MOC2 vs. MOC3: Z=–0.338, *p*=0.735). However, none of the pairwise comparisons would remain as statistically significant after Bonferroni correction for multiple comparisons [i.e., none of the p values was smaller than the corrected critical p<0.0083 (=0.05/6)].

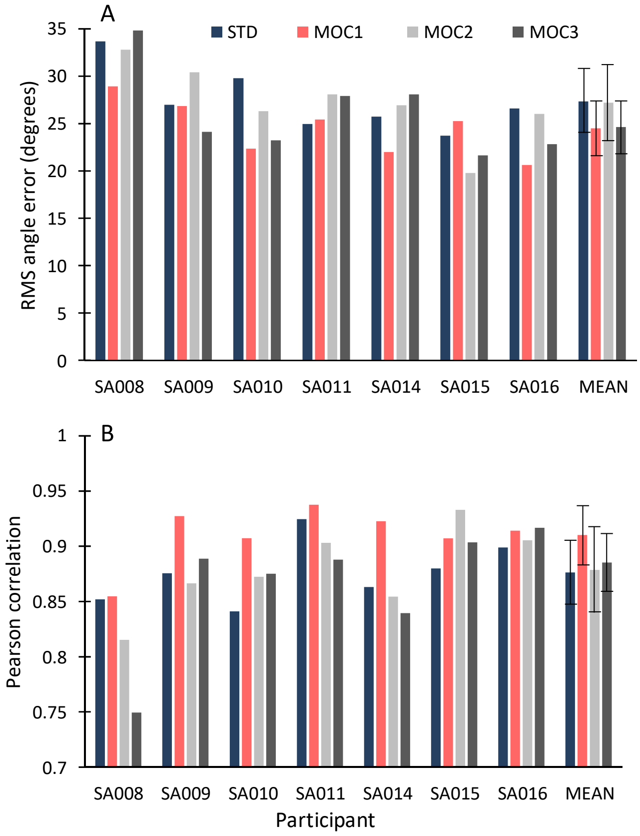

Figure. 5.9. Angle error (**A**) and Pearson correlation (**B**) between presentation and response angles for the STD, MOC1, MOC2 and MOC3 strategies. Values are shown for individual participants and the mean across participants. Error bars for the mean scores illustrate one standard deviation. MOC3 scores for participant SA008 are not included in the mean scores for the MOC3 strategy because the MOC3 strategy for this participant was implemented with the wrong BW_{ref} (see main text for details). Note that the scale in the ordinate of the bottom panel starts at 0.7 rather than zero to better show smaller differences. Individual scores for the STD and MOC1 strategies are re-plotted from **Figs. 5.6** and **5.7**.

5.4. DISCUSSION

We have shown that, compared to the STD strategy, the MOC1 strategy (with fast, binaurally coupled dynamic compression) enhances ILD cues (**Figs. 5.3** and **5.4**) and improves the localization of acoustic stimuli in a virtual horizontal plane (**Figs. 5.6** and **5.7**). Alternative implementations of the MOC1 strategy with slower (longer) time constants of integration (MOC2) and with greater inhibition in the lower than in the higher frequency channels (MOC3) can also enhance ILD cues for sufficiently long stimuli (**Fig. 5.2**). However, these more realistic implementations of the MOC strategy did not enhance the ILDs (**Fig. 5.4**) and did not improve the localization of the (short) 200-ms noise bursts used here relative to the STD strategy (**Fig. 5.9**).

5.4.1. Interpretation

Bilateral CI users rely mostly on ILD cues to judge the location of sound sources in the horizontal plane (Dorman et al., 2016; Laback et al., 2004; Litovsky and Gordon, 2016; Seeber and Fastl, 2008). Consistent with this, many aspects of the present results appear to be explained by the ILD versus azimuth functions produced by the tested strategies (**Fig. 5.4K**). For example, response matrices tended to flatten from azimuths of ±60° (**Fig. 5.5**) possibly because all strategies produced roughly constant ILDs for sound sources at and beyond 60°. Root-mean-square angle errors tended to be smaller at 60° (**Fig. 5.8A**) possibly because stimuli (and response screens) were bounded at 75°/90° and ILDs were approximately constant for azimuths ≥60°, leading listeners to response at 60°.

For the present (200-ms long) stimuli, the MOC1 strategy produced the largest ILDs and the steepest ILD-versus-azimuth function. Furthermore, ILDs for the MOC2, MOC3 and STD strategies were smaller than for the MOC1 strategy, and the corresponding ILD-versus-azimuth functions were shallower (**Fig. 5.4K**). This suggests that localization was overall better with the MOC1 strategy because this strategy provided ILDs that were larger than the just noticeable difference (JND) in ILD for the participants and coded for azimuth less ambiguously than any other strategy did.

The interaction between RMS angle error and azimuth was not statistically significant. Nonetheless, the localization improvements with the MOC1 strategy (re STD) tended to be larger in the frontal region (azimuths between –15° to 15°; **Fig. 5.8**), which is also the area with the smaller ILDs (**Fig. 5.4K**). This may be because at small angles, the ILDs provided by the STD strategy were smaller than or close to the JND-ILD of the listeners and became discernible with the MOC1 strategy (note that the JND-ILD is smaller at small angles; e.g., Fig. 3 in Yost and Dye, 1988).

5.4.2. Limitations

In measuring sound localization performance, it is common practice to rove the level of the acoustic stimulus to maximize the chance that localization be based on a ‘true’ interaural level cue rather than on the absolute level at either ear (e.g., Seeber et al., 2004; Majdak et al., 2011). Here, we roved the stimulus level for the three participants tested in North Carolina only but not for the nine participants tested in Salamanca. It is unlikely, however, that conclusions would have been different if the level had been roved for all participants. First, the monaural level-versus-azimuth functions at either ear were shallower than the corresponding ILD-versus-azimuth functions (**Fig. 5.4**). For example, for the MOC1 strategy and a stimulus level of –20 dB FS, the level at any ear changed by less than 10 dB over the –60° to 60° azimuth range while the corresponding ILD change was about 20 dB. This held true over a stimulus range from –40 to –10 dB FS. This indicates that the ILD was a more salient and possibly less ambiguous localization cue than the level at any single ear, even with roving of the stimulus level. Second, the trends in the data for the three participants tested with level roving was similar as for the other participants or the mean (e.g., angle errors were smaller, and correlations were greater with the MOC1 than with the STD strategy). However, monaural level cues might have been sufficient for localization if the level change across azimuths exceeded the level JND of the listener, particularly for the MOC1 strategy because it produced the steeper level-versus-azimuth functions (**Fig. 5.4**). Therefore, it cannot be entirely ruled out that the task could be performed to some uncertain extent by monitoring the stimulus level at a single ear.

The stimulus duration (200 ms) was shorter than the time required for a full activation (and deactivation) of contralateral inhibition in the MOC2 or MOC3 strategies (**Fig. 5.2**). As a result, the ILDs for the present stimuli were probably smaller than they would have been for longer stimuli. Indeed, vocoder simulations (not shown) revealed that the overall ILD for azimuth angles of ±60° would have been about 3 dB larger for the MOC2 strategy and about 1 dB larger for the MOC3 strategy if the stimulus duration had been 2 seconds rather than 200 ms (note that the use of longer stimuli would hardly increase the ILDs produced by the MOC1 strategy because contralateral inhibition was very fast in this strategy; i.e., τ_a and τ_b were both equal to 2 ms). Therefore, it is possible that the use of longer stimuli might improve localization performance with the MOC2 and MOC3 strategies to some uncertain extent.

Many individual participants showed better localization with the MOC1 than with the STD strategy (**Figs. 5.6** and **5.7**) even though the STD strategy was the most similar to the audio processing strategies worn by the participants in their clinical devices and participants were not given much opportunity to become fully accustomed to MOC processing before testing. **Figure 5.10** compares angle error scores across the practice session and the data collection session for those participants who had the two sessions. Error scores tended to be smaller in session 2 than in session 1 (i.e., most data points are below the diagonal), suggesting that performance tended to improve with practice. In addition, the vertical offset from the diagonal tended to be larger for participants with larger angle errors in the practice session, suggesting that practice benefited those who performed worse in the first session more than those who performed well. Last, the difference in performance across the two sessions (i.e., the potential effect of practice) tended to be smaller for the STD strategy than for any of the MOC strategies, possibly because the STD strategy provided localization cues most similar to those provided by the participants' own clinical devices. Altogether, this suggests that the potential benefits from the MOC1 strategy (and MOC processing in general) could become larger with practice and/or a sustained use of the MOC strategies.

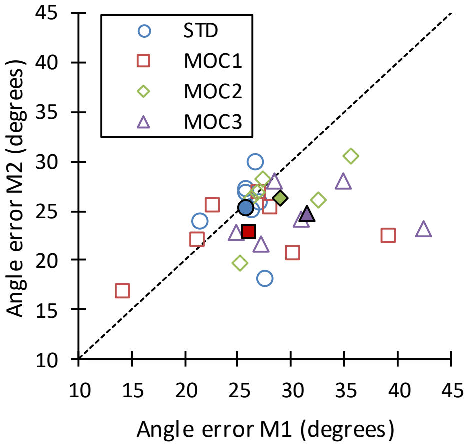

Figure 5.10. Comparison of angle error scores for the practice measurement session (M1, abscissa) and the test session (M2, ordinate). Different symbols illustrate results for different processing strategies, as shown by the inset. Open symbols illustrate results for six (MOC2 and MOC3 strategies) or seven (STD and MOC1 strategies) individual participants and filled symbols illustrate group mean results.

5.4.3. Comparison with related studies

The present tests were in simulated free-field conditions and the processing strategy used as the reference (the STD strategy) may have differed from the processing employed by the participants in their clinical devices. One might wonder (1) to what extent are the present results representative of lateralization in the free field? And (2) to what extent may the present findings generalize to clinical audio devices?

For listeners with normal hearing, the use of non-individualized HRTFs degrades the spectral details responsible for determining sound source elevation (e.g., Marmel et al., 2018) but to a large extent preserves the interaural difference cues responsible for determining the location of a sound source in the horizontal plane (Wenzel et al., 1993). Francart et al. (2011) reported that for listeners wearing a CI in one ear and a hearing aid in the other ear, the mean angle error for lateralization in a virtual sound field (28.4°) did not differ from that in the real sound field (31.5°). The angle errors obtained here with the reference, STD strategy (18° to 33°, mean = 25.3°, **Fig. 5.6**) were within the range of values reported in the literature for BiCI users tested with their clinical audio processors and for sound sources in the free field spanning a (broad) azimuth range similar to the one used here [e.g., the mean angle error in the free field was 24.5° in Nopp et al. (2004); 24.0° in Verschuur et al., (2005); 24.1° for noise and 21.5° for speech signals in Grantham et al., (2007); 20.4° for a wideband signal, 19.6° for a high pass signal, and 43.4° for a lowpass signal in Dorman et al. (2014); or 29.0° in Dorman et al. (2016)]. While some studies have reported smaller angle errors [e.g., 10° in van Hoesel and Tyler (2003)] or BiCI users performing close to normal (e.g., Seeber et al., 2004; Seeber and Fastl, 2008), this was probably due to using a narrower azimuth range over which the ILD-vs-angle is monotonic. For example, van Hoesel and Tyler (2003) used eight loudspeakers spaced at 15.5° and spanning 108° in front of the participant, and Seeber et al. (2004) and Seeber and Fastl (2008) used 11 speakers spaced at 10° from –50° to 50°. Therefore, altogether it is unlikely that the use of non-individualized HRTFs affected localization significantly. Even if it did, the effects of using non-individualized HRTFs would have been comparable across processing strategies. Altogether, this supports the conclusion that (1) the present results for the STD strategy are likely representative of the results that would be obtained with clinical devices in the free field; and (2) that it would not be unreasonable to generalize the reported effects of processing strategy to free-field tests.

We note, however, that the present tests were conducted without a front-end AGC. This differs from most clinical audio processors, which include a front-end broadband AGC compression stage (Zeng, 2004). In addition, we balanced the volume at the two ears to ensure that sentences at the two ears were perceived equally loud. This differs from typical clinical practice, where the output volume of each processor is set independently (Ching et al., 2007; Tyler et al., 2006). In addition, the HRTFs employed here were for a KEMAR (thus non-individualized) and were recorded with microphones placed at the eardrum position in a minimally reverberant (anechoic) room (Gardner and Martin, 1995). Therefore, the present HRTFs almost certainly provided different localization cues (input ILDs were possibly larger) than the participants were used to with their clinical audio processors in realistic, reverberant listening conditions. While binaural loudness balancing would seem appropriate in clinical practice and excluding AGC seems reasonable for isolating the effects of back-end compression on localization, participants may have adapted to different ILD-to-angle functions than they were used to with their devices in daily life.

Bilateral CI users lateralized more accurately with the MOC1 than with the STD strategy (mean angle error was 22.7° versus 25.3°, respectively). The MOC strategies were designed to reinstate the contralateral, dynamic control of compression mediated by the natural contralateral MOC reflex, which is absent for BiCI users (Lopez-Poveda et al., 2016b; Lopez-Poveda, 2018). If successful, one would expect the performance of BiCI users with the MOC strategies to be closer to the performance of listeners with normal hearing in the same task. The comparison remains to be done. It is unlikely, however, that BiCI users would show normal localization accuracy with the MOC1 strategy in realistic free-field settings. In natural listening environments, normal-hearing listeners have access to individualized ITD, ILD and spectral cues that would still be absent to BiCI users with the MOC strategy. Dorman et al. (2016) reported that in a free-field localization task with stimuli identical to the stimuli employed here and with a similar speaker arrangement, mean angle error scores were significantly greater (worse) for BiCI users than for young, normal-hearing listeners (29° versus 6°) and even the 'best' BiCI users had error scores above the 95th percentile of scores for young, normal-hearing listeners. This suggests that the mean angle error improvement of 2.6° provided by the 'best' MOC1 strategy would be insufficient to bring the performance of BiCI users equal to that of normal-hearing listeners, even if BiCI users were given sufficient practice on the MOC1 strategy.

5.5. CONCLUSIONS

(1) Compared to a STD strategy involving two independently functioning sound processors with fixed back-end compression, the MOC strategy with fast control of compression and greater inhibition in the higher frequency channels (MOC1), slightly improved the localization of wideband (125-6000 Hz) noise bursts in a virtual horizontal plane.

(2) MOC implementations that involved slower control of compression, and/or slightly greater inhibition in the lower than in the higher frequency channels (MOC2 and MOC3 strategies) also provided larger ILDs that the STD strategy for sufficiently long stimuli (>1 s). However, for the shorter (200-ms) noise bursts employed here, the localization performance with these strategies was not significantly different from that with the STD strategy.

(3) The localization improvements observed for the MOC1 strategy are probably due to this strategy providing larger and less ambiguous ILDs.

6.

SPEECH INTELLIGIBILITY WITH COMBINED MOC AND FS4 PROCESSING

6.1. INTRODUCTION

All evaluations of the MOC strategy conducted so far [including those reported in Chapters 4 and 5 as well as those reported by Lopez-Poveda et al. (2016b, 2017)] were restricted to a single target level of –20 dB FS, which corresponds approximately to 70 dB SPL in MED-EL clinical CIs. In addition, all tested MOC processors (MOC1, MOC2, and MOC3) and their corresponding standard reference processor (STD) were implemented without AGC and disregarded fine-structure (FS4) processing, even though AGC and FS4 processing are now standard in current MED-EL CI devices. To accurately assess the potential of MOC processing in an eventual implementation of the strategy in commercial MED-EL devices, it would be appropriate to compare speech-in-noise intelligibility with and without MOC processing in combination with AGC and FS4 processing, for a wider range of stimulus levels and types of maskers. This chapter reports a study aimed at making this comparison.

Intelligibility in noise was compared for a standard FS4 processing strategy (STD-FS4) and for an FS4 strategy with MOC3-type contralateral control of back-end compression (MOC3-FS4). The two strategies were implemented with identical 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, it could be easily implemented in a binaural CI device and would theoretically preserve head-shadow ILDs that are useful for contralateral MOC processing to function properly. The two strategies (STD-FS4 and MOC3-FS4) have been described in Section 3.2.

Speech reception thresholds were measured with the two strategies for sentences in quiet and in competition with a single-source of steady-state SSN or an international female fluctuating masker (iFFM) (Holube et al., 2010). For the SSN masker, SRTs were measured for speech at three different levels (–48, –38, and –28 dB FS). For the iFFM, SRTs were measured for speech at –38 dB FS. These levels were chosen because the default level of –38 dB FS with AGC corresponds to about –20 dB FS without AGC, the level tested in previous MOC studies that corresponds to a typical conversational level of 70 dB SPL. Speech reception thresholds were measured in bilateral and unilateral listening (with the self-reported better ear) and for multiple spatial configurations of the target and masker sources (described below).

6.2. METHODS

6.2.1. *Participants*

Seven BiCI users participated in the experiments (**Table 3.2**). Three of them were teenagers (SA025, SA026 and SA027), and four were adults (SA021, SA022, SA023 and SA024). Except for SA022, who had participated in previous tests of the MOC strategy (identified as SA014 in Chapter 4), participants had not been tested in the laboratory before.

6.2.2. *Stimuli*

Speech reception thresholds for sentences in quiet and in noise were measured using the procedures described in Section 3.4. To reduce the potential confounding effects of participants learning the test sentences, an attempt was made to measure SRTs using the female sentences in the Spanish version of the Oldenburg Sentence Test (or matrix test) for all participants (Hochmuth et al., 2012). However, some participants found it impossible to understand 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 (Huarte, 2008) of the HINT (Nilsson et al., 1994). The sentence material used to test each participant is shown in **Table 6.1**. The use of different sentence material for different participants was deemed reasonable because the aim was to compare performance across two processing strategies tested with the same speech material, rather than to compare performance across participants. In other words, any effect of speech material was assumed to affect the two strategies equally.

Speech reception thresholds were measured for sentences masked by SSN and an iFFM. The SSN from the matrix test or the HINT was used when SRTs were measured using matrix and HINT sentences, respectively. 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. Speech reception thresholds were measure using fixed-level speech and varying the noise level adaptively. 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.

Table 6.1. 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						
SA021	*Speech* *Noise*	HINT SSN	HINT SSN	HINT SSN	n.m.	HINT iFFM
SA022	*Speech* *Noise*	Matrix Matrix noise	Matrix Matrix noise	Matrix Matrix noise	Matrix Matrix noise	Matrix iFFM
SA023	*Speech* *Noise*	Matrix Matrix noise	Matrix Matrix noise	Matrix Matrix noise	Matrix Matrix noise	Matrix iFFM
SA024	*Speech* *Noise*	HINT SSN	HINT SSN	HINT SSN	HINT SSN	HINT iFFM
SA025	*Speech* *Noise*	Matrix Matrix noise	Matrix Matrix noise	n.m.	n.m.	Matrix iFFM
SA026	*Speech* *Noise*	Matrix Matrix noise	Matrix Matrix noise	Matrix Matrix noise	n.m.	Matrix iFFM
SA027	*Speech* *Noise*	HINT SSN	HINT SSN	HINT SSN	HINT SSN	HINT iFFM

6.2.3. 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 (**Table 3.2**). 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 as well as 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 $S_{90}N_{-90}$).

Speech reception thresholds were measured for the following spatial configurations:

- S_0N_0: with the speech and masker sound sources co-located in front of the listener at 0° azimuth.
- $S_{15}N_{-15}$: with the speech and masker sound sources at 15° and –15°, respectively, when the self-reported better ear was the right ear, or –15° and 15° when the better ear was the left ear. The nomenclature $S_{15}N_{-15}$ will be used for convenience.
- $S_{60}N_{-60}$: with the speech and masker sound sources at 60° and –60°, respectively, when the self-reported better ear was the right ear, or –60° and 60° when the better ear was the left ear. The nomenclature $S_{60}N_{-60}$ will be used for convenience.
- $S_{90}N_{-90}$: with the speech and masker sound sources at 90° and –90°, respectively, when the self-reported better ear was the right ear, or –90° and 90° when the better ear was the left ear. The nomenclature $S_{90}N_{-90}$ will be used for convenience.
- S_0N_{-90}: with the speech and masker sound sources at 0° and –90°, respectively, when the self-reported better ear was the right ear, or 0 and 90° when the better ear was the left ear

(i.e., the speech source was in front with the masker source contralateral the self-reported better ear). The nomenclature S_0N_{-90} will be used for convenience.

As a control (see below), SRTs were measured also in quiet for a sound source at 0° azimuth. To minimize re-using the test sentence lists, SRTs in quiet were measured using the practice sentences from the HINT and matrix tests, and only one SRT estimate was obtained per condition. Speech reception thresholds in quiet were measured in unilateral listening with the left and right CIs, and in bilateral listening.

6.2.4. Processing strategies

Speech reception thresholds were measured for stimuli processed through the STD-FS4 and MOC3-FS4 strategies, which have been described in Chapter 3.

6.2.5. Order of testing

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

1. Bilateral listening with speech level at –38 dB FS and with the SSN masker.
2. Bilateral listening with speech level at –38 dB FS and with the iFFM masker.
3. Bilateral listening with speech level at –28 dB FS and with the SSN masker.
4. Bilateral listening with speech level at –48 dB FS and with the SSN masker.
5. 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, 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). 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 6.1**).

As a control (see below), SRTs in quiet were obtained prior to measuring the SRTs in noise. SRTs in quiet were measured for the STD-FS4 and MOC3-FS4 strategies, in bilateral listening and in unilateral listening with both the left and right ears. The target source was at 0° azimuth. Only one SRT estimate was obtained per strategy per condition.

6.3. RESULTS

6.3.1. Speech reception thresholds in quiet

The aim of the present study was to compare SRTs in noise for the STD-FS4 and MOC3-FS4 strategies, in bilateral and unilateral listening and for speech levels as low as –48 dB FS. Implicit in the experimental design, there were three assumptions:

(1) That all participants would be able to recognize sentences presented at –48 dB FS, the lowest speech level tested, i.e., that audibility would not be an issue for discriminating speech at –48 dB FS. Note that speech discrimination could be difficult at low levels, particularly with the MOC3-FS4 strategy because it can theoretically reduce audibility due to contralateral inhibition (see Chapter 4).
(2) That participants were accurate in reporting their better ear, i.e., that SRTs would be lower when listening with the self-reported better than with the worse ear.
(3) That the performance of BiCI users in bilateral listening conditions was *not* dominated by the self-reported better ear, i.e., that they obtained some benefit from listening with two ears.

We tested these assumptions by measuring SRTs in quiet in bilateral listening as well as in unilateral listening with the left and right ears. The results from these pilot tests (shown in **Fig. 6.1**) revealed the following:

1. Almost all SRTs in quiet were lower (better) than –38 dB FS, the default speech level used in the main tests. The exception was the SRT for the STD-FS4 strategy for the right ear (the worse ear) of participant SA021 (i.e., the SRT in question is above the dashed line at –38 dB FS in **Fig. 6.1**). In addition, almost all SRTs in quiet were better than –48 dB FS, the lowest speech level used in the main tests. The exception was the right ear (the worse ear) of participant SA021, whose SRTs are above the dotted line at –48 dB FS. This confirms that, with one exception, participants were able to understand speech at the lowest levels tested in the main experiment when listening bilaterally or unilaterally with either ear.

2. For the STD-FS4 strategy (the closer to the clinical strategy), SRTs were equal or lower for the self-reported better ear than for the worse ear of all participants (i.e., open symbols overlapped with or were below the filled symbols in **Fig. 6.1**). This confirms that all participants were accurate in reporting their better ear.

3. For the STD-FS4 strategy, SRTs were equal or only slightly better when listening bilaterally than with the self-reported better ear alone (i.e., triangles overlapped with or were below the open symbols in **Fig. 6.1**). The exception was participant SA024, which performed worse when listening bilaterally than with the better ear alone. Altogether, these data suggest that participants obtained little benefit from listening in quiet with two implants.

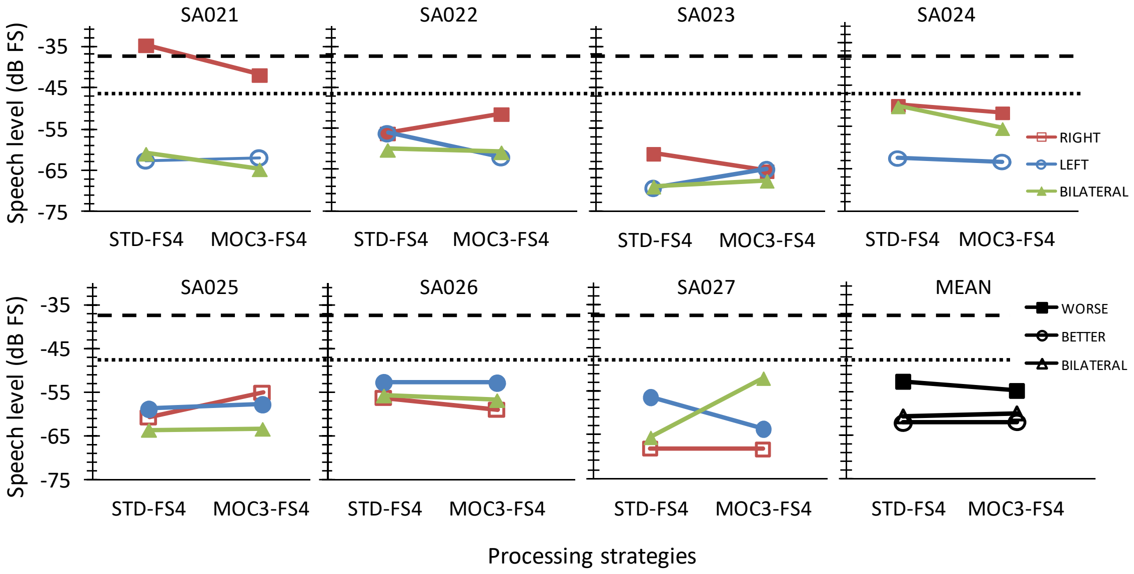

Figure 6.1. Individual SRTs (in dB FS) for sentences in quiet for the STD-FS4 and MOC3-FS4 strategies. Each panel shows values for a single participant or the mean across participants, as indicated at the top of the panel. Different symbols illustrate SRTs for different listening conditions: listening bilaterally (triangles); listening with the right ear alone (squares); and listening with the left ear alone (circles). For panels showing individual data, open symbols depict individual scores for the self-reported better ear (**Table 3.2**). The horizontal dashed and dotted lines depict the default (–38 dB FS) and the lowest (–48 dB FS) speech level used in the main tests, respectively.

6.3.2. Benefits from MOC3-FS4 processing in bilateral listening

Speech level at –38 dB FS in SSN and iFFM maskers

The top row in **Fig. 6.2** (panels **A**-**E**) show SRTs in bilateral listening with the STD-FS4 strategy for speech at –38 dB FS in competition with SSN (recall that each point is the mean of three estimates). Blue and red bars show individual and group mean (M) scores, respectively. Each panel is for a different spatial configuration (S_0N_{90}, S_0N_0, $S_{15}N_{-15}$, $S_{60}N_{-60}$ and $S_{90}N_{-90}$), as indicated at the top of the panel. The bottom row (panels **F**-**J**) illustrates SRT improvements (in dB, Eq. 4.1) 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.

The figure shows that the MOC3-FS4 strategy tended to be advantageous on average over the STD-FS4 strategy for all spatial configurations except for the co-located S_0N_0 condition (**Fig. 6.2G**), where SRTs were approximately equal for the STD-FS4 and MOC3-FS4 strategies. Friedman's test (for related samples and related groups) revealed a significant main effect of processing strategy on SRTs (i.e., when SRTs across spatial configurations were grouped together) [$\chi^2(1)=8.25$, p=0.004]. Friedman's test also showed a significant effect of test condition (as defined by processing strategy and spatial configuration) on SRTs [$\chi^2(9)=44.58$, p<0.001]. Post-hoc pairwise Wilcoxon signed-rank tests without Bonferroni corrections showed a significant difference in SRTs for the STD-FS4 and MOC3-FS4 at $S_{15}N_{-15}$ (p=0.028) and at $S_{90}N_{-90}$ (p=0.043). However, when the *p* values were corrected for multiple comparisons, the effect of strategy did not remain as significant for any spatial configuration.

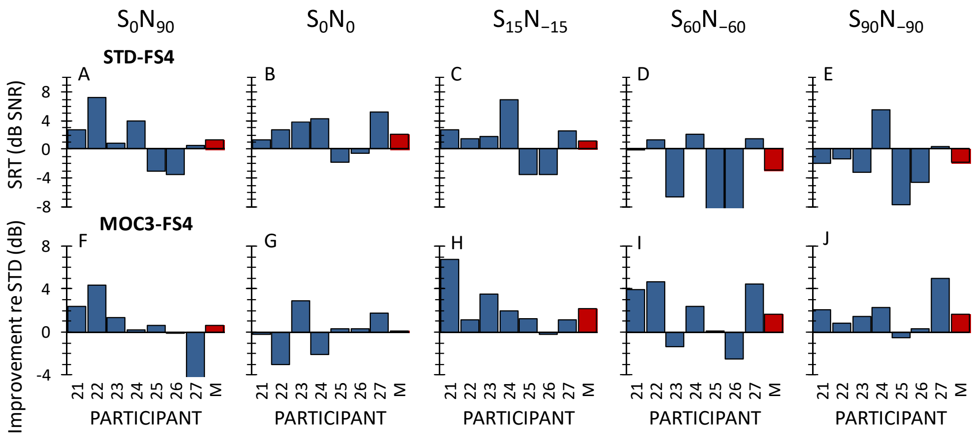

Figure 6.2. Top (A-E). Speech reception thresholds in bilateral listening with the STD-FS4 strategy for speech at −38 dB FS and SSN masker. Each panel illustrates individual and group mean SRTs. Each panel is for a different spatial configuration, as indicated at the top. **Bottom (F-J)**. Speech reception threshold improvement provided by the MOC3-FS4 strategy relative to the STD-FS4 strategy. Blue and red bars show individual and group mean (M) scores. See main text for details.

For the iFFM masker (**Fig. 6.3**), the MOC3-FS4 strategy seemed slightly advantageous on average only for the $S_{90}N_{-90}$ spatial configuration (**Fig. 6.3J**) and was not different from the STD-FS4 in any of the other spatial configurations tested. Friedman's test (for related samples and related groups) revealed that the main effect of processing strategy on SRTs (with SRTs across spatial configurations grouped together) was not significant [$\chi^2(1)=0.714$, $p=0.398$]. Friedman's test revealed a significant effect of test condition (given by processing strategy and spatial configuration) [$\chi^2(9)=54.82$, $p<0.001$]. However, post-hoc pairwise Wilcoxon signed-rank tests with and without Bonferroni corrections did not show a significant effect of processing strategy for any of the spatial configurations tested.

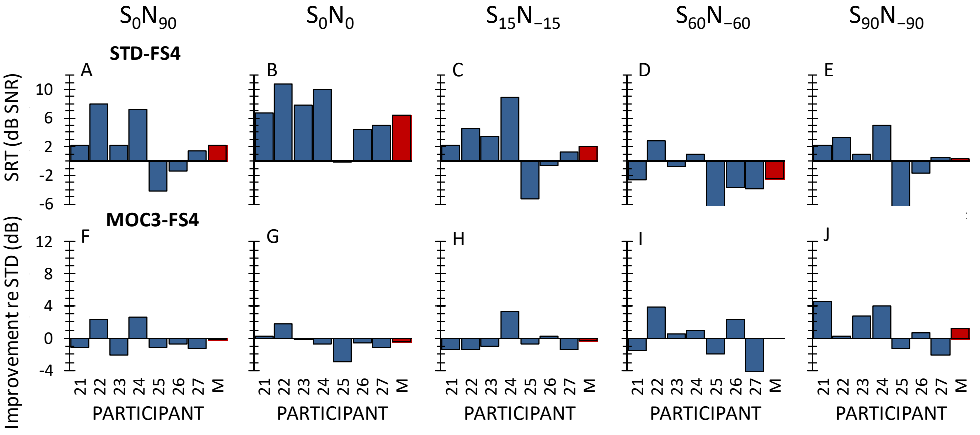

Figure 6.3. The same as **Fig. 6.2** but for an iFFM masker.

Speech level at −28 and −48 dB FS in SSN masker

For speech at −28 dB FS in competition with SSN, SRTs were equal or better with the MOC3-FS4 strategy than with the STD-FS4 strategy for all spatial configurations (**Fig. 6.4**). Friedman's test revealed a significant main effect of strategy on SRTs (i.e., when SRTs across spatial configurations were grouped together) [$\chi^2(1)=5.76$, $p=0.016$]. Friedman's test also revealed a significant effect of test condition (as defined by processing strategy and spatial configuration) [$\chi^2(5)=27.16$, $p<0.001$]. However, post-hoc pairwise Wilcoxon signed-rank tests without Bonferroni corrections showed that the SRTs for the STD-FS4 and MOC3-FS4 were not significantly different for any of the spatial configurations tested.

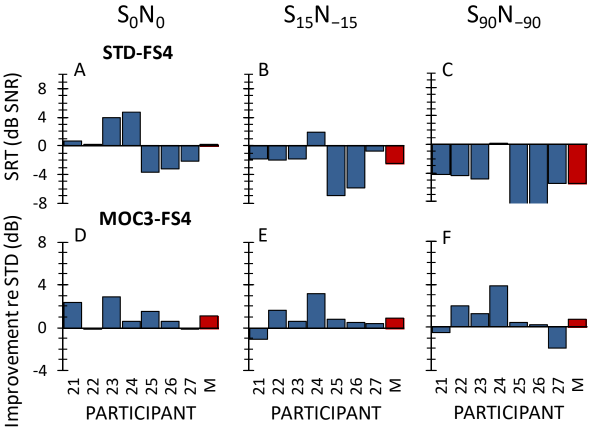

Figure 6.4. Speech reception thresholds in bilateral listening for speech at –28 dB FS in competition with SSN and for three different target-masker spatial configurations (as indicated at the top of each panel). The layout is the same as for **Fig. 6.2**.

For speech at –48 dB FS in competition with SSN, SRTs were equal or better with the MOC3-FS4 strategy than with the STD-FS4 strategy for all spatial configurations (**Fig. 6.5**). Friedman's test showed that the main effect of strategy on SRTs was not significant [$\chi^2(1)=0.529$, $p=0.467$]. Friedman's test revealed that SRTs were significantly differences across test conditions [$\chi^2(5)=21.65$, $p=0.001$]. However, pairwise post-hoc Wilcoxon signed-rank tests with and without Bonferroni corrections did not show significant effects of processing strategy for any of the spatial configurations tested.

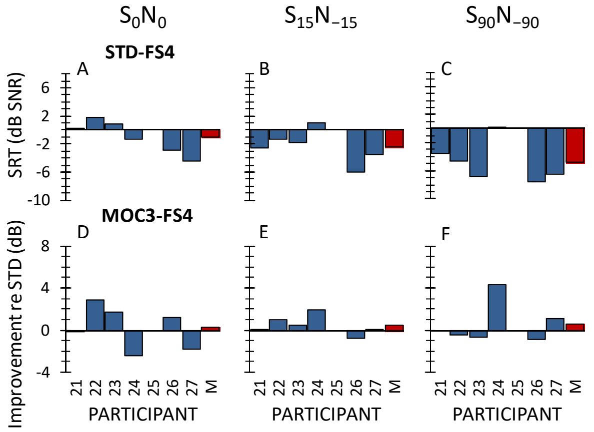

Figure 6.5. As **Fig. 6.4** but for a speech level of –48 dB FS.

6.3.3. Benefits from MOC3-FS4 processing in unilateral listening

In unilateral listening, SRTs were measured 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. 6.6**. On average, SRTs were equal or better with the MOC3-FS4 than with the STD-FS4 strategy for all spatial configurations. Friedman's test for related samples and related groups revealed a statistically significant main effect of strategy [$\chi^2(1)=8.33$, p=0.004]. Friedman's test also showed a significant effect of test condition [$\chi^2(7)=25.50$, p=0.001]. However, post-hoc Wilcoxon signed-rank tests with and without Bonferroni corrections did not show significantly different SRTs for STD-FS4 and MOC3-FS4 strategies for any of the spatial configurations tested.

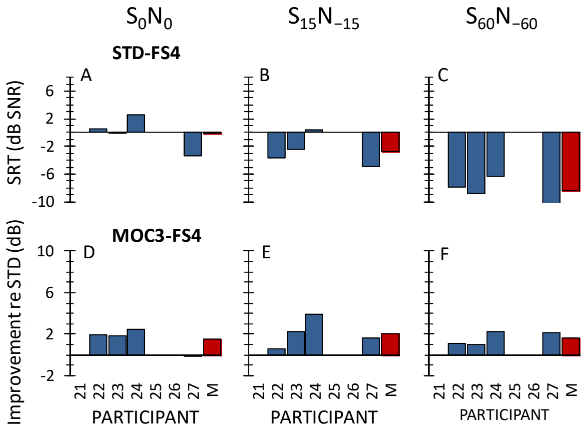

Figure 6.6. Speech reception thresholds in unilateral listening for speech at –38 dB FS in competition with SSN, and for three different target-masker spatial configurations (as indicated at the top of each panel). The layout is the same as for **Fig. 6.2**.

6.3.4. Speech-in-noise recognition in bilateral versus unilateral listening

In this section, we compare SRTs in noise measured in bilateral versus unilateral listening. Recall that unilateral listening involved listening with the self-reported better ear. Therefore, this section addresses the question: in noisy environments, was listening with two CIs advantageous over listening with the better-ear CI alone? In other words, was there a binaural advantage?

Figure 6.7 shows mean SRTs in SSN noise for speech at –38 dB FS for the STD-FS4 and MOC3-FS4 strategies in unilateral (circles) and bilateral listening (triangles). For all participants and conditions, SRTs were always equal or lower in unilateral than in bilateral listening. This result was unexpected and possible explanations are discussed below.

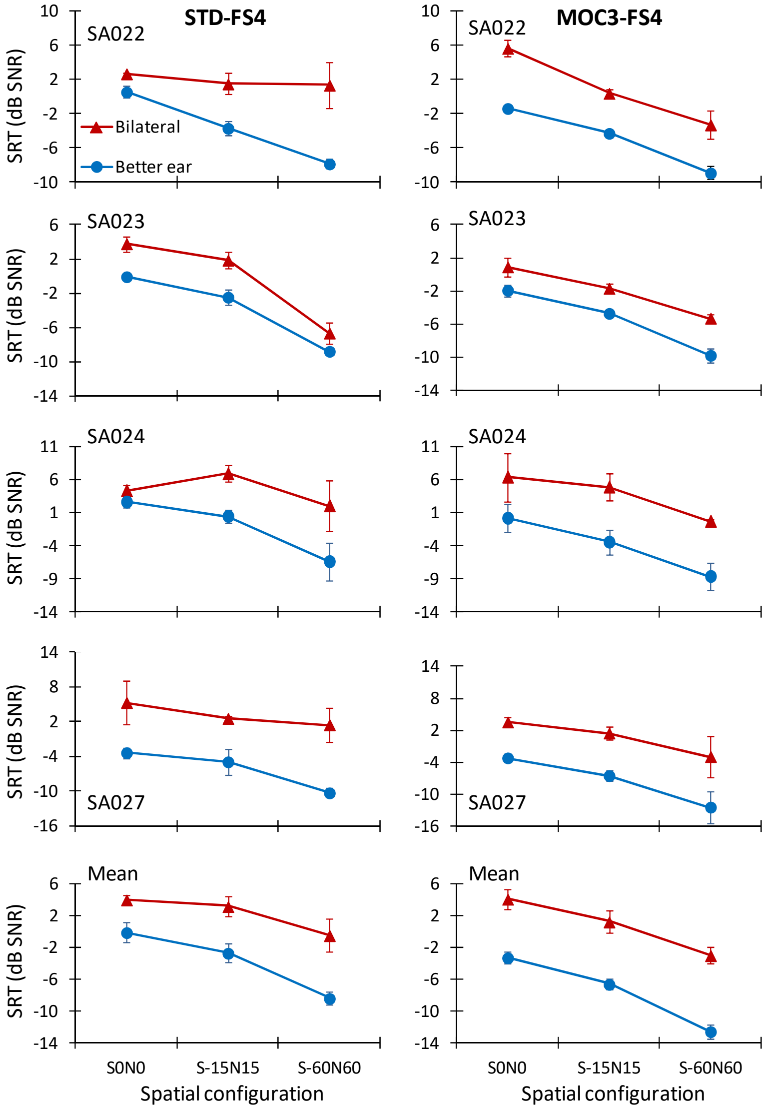

Figure 6.7. Mean SRT in noise as a function of target-masker spatial configuration for the STD-FS4 (**left column**) and MOC3-FS4 (**right column**) strategies in unilateral (circles) and bilateral listening (triangles). The speech level was fixed at –38 dB FS and the masker was SSN. Each row is for an individual participant, as indicated at the top of each panel. The bottom row shows mean SRTs across four participants (SA022, SA023, SA024 and SA027). For individual participant data, each point is the mean of three SRT estimates. Error bars indicate one standard error of the mean (N=3).

6.4. DISCUSSION

6.4.1. Why were the benefits of MOC3-FS4 greater in SSN than in an iFFM?

For speech at –38 dB FS, the benefits of the MOC3-FS4 strategy (re STD-FS4) tended to be greater for speech in competition with a stationary (SSN) than with a fluctuating (iFFM) masker. For the SSN masker, MOC3-FS4 improved average SRTs (re STD-FS4) for all spatial configurations except S_0N_0 (**Fig. 6.2F-J**), while the benefit with the iFFM occurred only for the $S_{90}N_{-90}$ spatial configuration (**Fig. 6.3F-J**).

The benefit of MOC3-FS4 processing relative to STD-FS4 might be overall greater for SSN than for iFFM because the masker output level (which was used to determine the *c* parameter in the MOC3-FS4 strategy) was (1) overall smaller for the iFFM than for SSN because the iFFM was fluctuating and had silent gaps; and (2) was more similar at the two ears for the iFFM than for SSN because with the iFFM, the two ears were effectively 'sensing' speech (sparse) signals. Hence, the effective contralateral inhibition was possibly less and more similar at the two ears for the iFFM than for SSN. The use of a faster temporal window of integration (with shorter time constants), as in the MOC1 strategy (Chapter 4), might increase contralateral inhibition with the iFFM because the corresponding output level would follow the more rapid variations in iFFM amplitude.

6.4.2. 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. 6.2, 6.4 and 6.5**). This is possibly because the MOC3 parameters were carried forward from the study described in Chapter 4, where they were optimized for a strategy without AGC or FS4 and for a speech level of –20 dB FS, which roughly corresponds to the –38 dB FS when AGC is used. Both these levels (–20 dB FS without AGC and –38 dB FS with AGC) correspond to 70 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 greater benefit for speech levels of about 70 dB SPL but would also provide some benefit (albeit smaller) for speech levels of 60 and 80 dB SPL.

6.4.3. Why was speech-in-noise recognition better in unilateral than in bilateral listening?

We have unexpectedly found that SRTs in noise were always better in unilateral than in bilateral listening (**Fig. 6.7**). The reason is uncertain. One possible factor is learning. Bilateral listening tests were conducted on the first day of testing while unilateral listening tests were conducted towards the end of a five-day test visit (see Section 6.2.5). That is, unilateral listening tests were administered typically after four days of testing and after up to 96 SRT measurements in various other bilateral listening conditions. Therefore, participants could have learnt to perform the task by the time unilateral tests were conducted.

Figure 6.8 supports the learning explanation by illustrating SRTs as a function of test number for six participants. Note that three participants were tested using matrix sentences while the three other participants were tested using HINT sentences. Speech reception thresholds are shown for all test conditions combined. Clearly, for these six participants, SRTs improved gradually (became lower) with increasing test number, even though unilateral listening tests (light blue symbols), which should presumably yield equal or higher (worse) SRTs if there was a binaural advantage, were conducted last and bilateral listening tests were conducted first (dark blue symbols).

Further evidence that learning occurred is that SRTs in bilateral listening and in SSN were higher for the first test condition (dark blue symbols) 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 with the lower speech level presumably because this was the fourth test condition and was administered towards the end of the week.

Also striking was that learning appeared to be similar for participants tested with HINT and matrix sentences, even though matrix sentences were specifically designed to minimize learning effects by omitting semantic information (Hochmuth et al., 2012). This suggests that the learning effect illustrated in **Fig. 6.8** was unlikely due to participants learning the sentences and instead, it was more likely due to participants learning how to perform the task. This finding should be considered when interpreting past and previous findings.

For each test condition [defined by the speech level, masker type (SSN or iFFM) and listening modality (bilateral listening or unilateral listening with the better ear)], the learning effects appeared to be similar to the learning effects along the whole set of data except for the first test condition, where learning was more obvious (the slope was steepest). On the other hand, the learning effects were only significant when comparing conditions that were measured well apart in time, such as bilateral (points on the far left of each panel) versus unilateral listening conditions (points on the far right of each panel). Because each testing block included SRT measurements that combined the two strategies and spatial configurations in random order, learning unlikely affected conclusions on the effect of strategy, spatial configuration, or their interaction on the SRTs measured within any one of the five tested conditions (see below). The same is true for previous MOC studies. Learning, however, likely rendered comparisons of SRTs across conditions (e.g., speech level or listening modality) unreliable because different conditions were tested on different days as explained above.

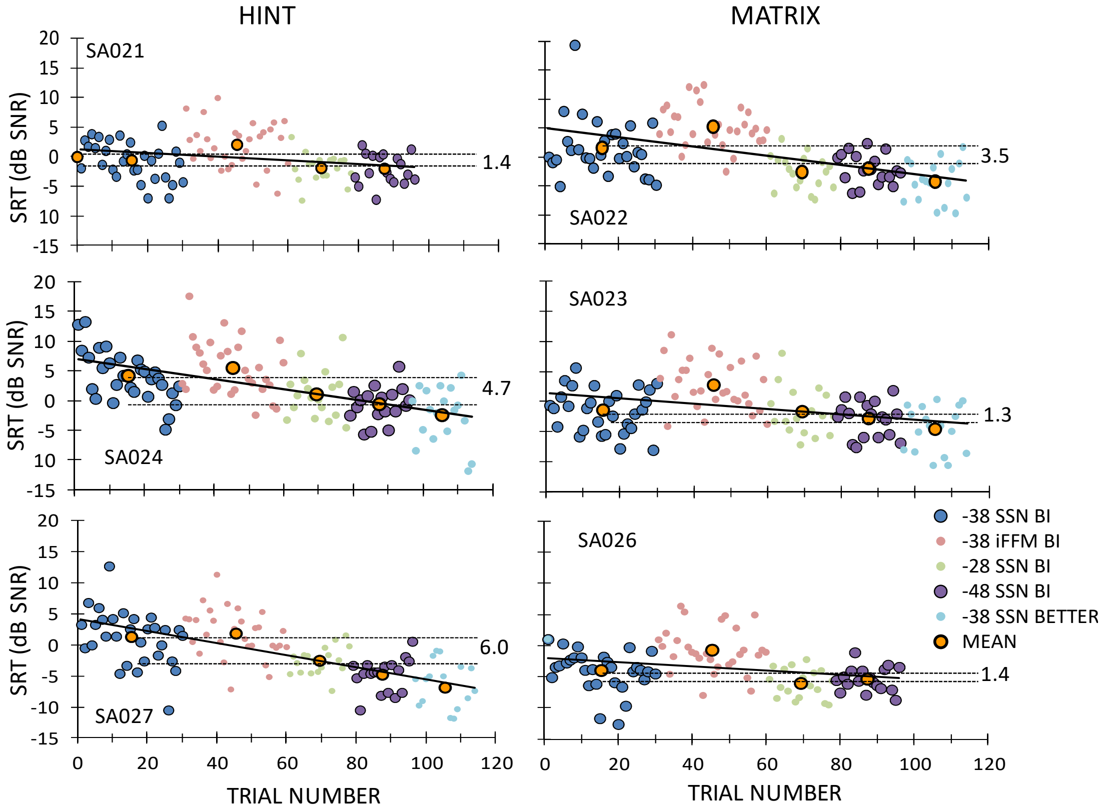

Figure 6.8. Speech reception thresholds as a function of test order for six participants (SA021, SA022, SA023, SA024, SA026 and SA027). Values are pooled for all tested conditions, including different sound-processing strategies, target-masker spatial configurations, masker type, and speech level. Different colors illustrate groups of SRTs for a different 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 (BETTER), 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 SA022, SA023 and SA026 were tested using matrix sentences, while participants SA021, SA024 and SA027 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.

Learning, however, needs not be the only factor that contributed to the better performance in unilateral than in bilateral listening. Some studies have shown that bilateral stimulation can sometimes produce worse speech recognition than unilateral stimulation (e.g., Goupell et al., 2018). One possible reason for this is related with stimulus-encoding factors. That is, the binaural processing of sounds might be stunted by interaural mismatch in the cochlear tonotopic place of stimulation (Xu et al., 2020). This means that the corresponding frequency regions of the sound reaching the two ears may not conduct electrodes innervating coincident regions of the two cochleae, thus producing signals likely distorted differently in the two ears (Long et al., 2006; Goupell and Litovsky, 2015; Goupell et al., 2015; Kan et al., 2015). A binaural mismatch can be caused by a device-related factor (the type of electrode array), a surgical factor (placement of the electrode array in the cochlea), or a biological factor (neural survival) (Goupell et al., 2018). Here, stimuli were processed with identical frequency maps for the two ears (**Table 3.2**), but electrode arrays, their placement in each participant's cochleae, and neural survival could have been mismatched across the ears. This mismatch could have decreased binaural sensitivity, reduced binaural fusion, decreased speech understanding, and distorted and decreased perceived lateralization for ITD and ILD cues (Goupell et al., 2013; Goupell et al., 2015; Kan et al., 2015; Goupell et al., 2018; Xu et al., 2020). Therefore, it is conceivable that bilateral stimulation might have been disadvantageous because of mismatched electrical stimulation across the ears for the present participants (Litovsky et al., 2012; Kan and Litovsky, 2015; Goupell et al., 2018).

Unlike the present study, in the study described in Chapter 4 we found listening with two CIs to be advantageous over listening with one CI (Section 4.3.3 and **Fig. 4.8**). In contrast to the present study, however, the study described in Chapter 4 involved sound-processing strategies without AGC or FS4 processing, and unilateral listening tests were conducted before (rather than after) bilateral listening tests. Further research would be necessary to elucidate the factor(s) causing the differences in binaural advantage across the two studies.

6.5. CONCLUSIONS

The aim of the present study was to compare speech-in-noise intelligibility with the MOC3-FS4 strategy against that with the STD-FS4 strategy. Speech reception thresholds were compared for sentences presented in competition with steady-state (SSN) and fluctuating (iFFM) maskers and for various speech levels (–28, –38, and –48 dB FS). The main findings were:

(1) For BiCI users tested in bilateral listening, mean SRTs were 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) (**Fig. 6.2** to **Fig. 6.5**). Across all listening conditions and spatial configurations, the mean SRT improvement re STD-FS4 was about 1 dB. The largest benefit was for speech at –38 dB FS and SSN (**Fig. 6.2**). In the latter condition, the mean SRT improvement re STD-FS4 across spatial locations was 1.2 dB.

(2) For bilateral CI users tested in unilateral listening, mean SRTs were equal or better with the MOC3-FS4 than with the STD-FS4 strategy. The mean SRT improvement across the tested target-masker spatial configurations was about 1.7 dB (**Fig. 6.6**).

(3) The benefits provided by the MOC3-FS4 strategy (re STD-FS4) tended to be greater for speech in competition with SSN than with iFFM.

(4) Speech reception thresholds were consistently worse in bilateral than in unilateral listening for both the STD-FS4 and MOC3-FS4 strategies. The reason is uncertain.

7.

LISTENING EFFORT WITH VARIOUS IMPLEMENTATIONS OF THE MOC STRATEGY

7.1. INTRODUCTION

In Chapters 4 and 6 as well as in previous studies (Lopez-Poveda et al., 2016a, 2017), it has been shown that the MOC strategy can improve the intelligibility of speech in noise for CI users. However, the effect of MOC processing on listening effort is yet to be investigated. This chapter aims at investigating whether the MOC strategy affects listening effort during speech recognition. The chapter integrates two studies conducted at different times. In Study 1, a dual-task method was used to compare listening effort with the STD, MOC1, MOC2, and MOC3 strategies for BiCI users. The primary task involved recognizing words in quiet and in noise, and the secondary task involved recalling the words. Because the MOC strategy can facilitate the recognition of speech in noise, we hypothesized that in noise and at any given SNR, the MOC strategy reduces listening effort compared to the STD strategy. We also hypothesized that in noise and for conditions of equal intelligibility (i.e., for speech at the individual SRT), listening with the MOC strategy requires the same effort as listening with the STD strategy, indicating that the better intelligibility in noise with the MOC than with the STD strategy is not the result of participants spending more effort with the MOC strategy.

In Study 2, we used the same dual-task paradigm as in Study 1 but we also measured the verbal response time to compare listening effort with the STD-FS4 strategy and with the MOC3-FS4 strategy (for more details see Chapter 3). Verbal response times were quantified as the time between the end of the stimulus presentation (a disyllabic word) and the participant’s response (repetition of the word). Response times were recorded for both correct and incorrect responses and were averaged across three trials for each participant. We hypothesized that response times would be equal or lower with the MOC3-FS4 as with the STD-FS4 strategies, indicating that listening with the MOC3-FS4 strategy requires equal or less effort than listening with the STD-FS4 strategy.

In what follows, the methods and discussion of the two studies are described in an integrated manner. For convenience, however, the results of the two studies are presented separately.

7.2. METHODS

7.2.1. Participants

Six bilateral CI users participated in Study 1. Two of them were teenagers (SA009 and SA010), and four were adults (SA011, SA014, SA015 and SA016). They had between 9 and 12 active electrodes in their implants (**Table 3.1**).

Seven BiCI users participated in Study 2 (**Table 3.2**). Three of them were teenagers (SA025, SA026 and SA027), and four were adults (SA021, SA022, SA023 and SA024). Only participant SA022 was common to the two studies (identified as SA014 in Study 1).

7.2.2. Stimuli

In Study 1, listening effort was assessed as described in Section 3.6.1. The target words were presented at a fixed level of –20 dB FS. In Study 2, listening effort was assessed as described in Sections 3.6.1 and 3.6.2. The words were presented at a fixed level of –38 dB FS. The latter level corresponds approximately to –20 dB FS once stimuli are passed through the linked AGCs involved in the STD-FS4 and MOC3-FS4 strategies tested in Study 2. In the two studies, the masker was SSN. All tests involved bilateral stimulation.

Word recognition was measured for three SNRs: in quiet, at +5 dB SNR, and at the individual SNR giving 50% correct sentence-in-noise recognition. We will refer to the latter as the individual SRT. We chose to measure effort in quiet because MOC processing can reduce audibility slightly, which might increase effort if the reduction in audibility degraded intelligibility. We chose +5 dB SNR because it is a typical SNR in natural listening situations (Smeds et al., 2015). We decided to measure effort at the individual SRT to investigate the effect of processing strategy on effort in conditions of equal intelligibility. Effort typically decreases with increasing intelligibility and for a fixed SNR, the MOC strategy can improve intelligibility in noise. Therefore, we expected MOC processing to decrease or not change listening effort for speech at a fixed SNR. However, we wished to demonstrate that MOC processing did not require greater effort than STD processing in conditions of equal intelligibility.

The individual SRT values used in Study 1 and Study 2 were obtained from the intelligibility studies described in Chapter 4 and Chapter 6, respectively, for bilateral listening conditions (**Fig. 4.6** and **Fig. 6.4**). The actual SRT values used in Study 1 and Study 2 are shown in **Table 7.1** and **Table 7.2**, respectively. 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 50 (Study 1) or 100 ms (Study 2) after the word offset.

Table 7.1. Speech perception thresholds in bilateral listening for sentences in SSN (in units of dB SNR) for each participant of Study 1, strategy (STD, MOC1, MOC2, and MOC3), 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. 4.6**.

		Strategy			
Condition	**Participant**	**STD**	**MOC1**	**MOC2**	**MOC3**
$S_{15}N_{-15}$	SA009	–3.3	–3.7	–4.7	–3.3
	SA010	–7.8	–7.9	–8.5	–8.1
	SA011	–4.6	–4.3	–4.9	–5.4
	SA014	–2.3	–4.8	–4.1	–4.9
	SA015	–5.9	–4.8	–7.5	–7.6
	SA016	–3.8	–4.9	–4.0	–4.9
	Mean	–4.0	–4.9	–5.2	–5.1
	s.d.	2.5	1.4	2.1	2.2
$S_{60}N_{-60}$	SA009	–5.3	–6.9	–7.1	–7.7
	SA010	–13.7	–15.9	–10.6	–15.2
	SA011	–7.3	–8.2	–7.0	–9.6
	SA014	–8.6	–9.5	–10.5	–8.4
	SA015	–10.8	–10.5	–11.1	–12.1
	SA016	–5.4	–9.1	–6.5	–7.0
	Mean	–8.3	–9.3	–8.2	–9.3
	s.d.	3.0	3.4	2.5	3.5

Table 7.2. As **Table 7.1** but for Study 2. Speech reception thresholds values correspond to those shown in **Fig. 6.2**.

		Strategy	
Condition	**Participant**	**STD-FS4**	**MOC3-FS4**
$S_{15}N_{-15}$	SA021	2.6	–4.0
	SA022	1.5	0.4
	SA023	1.8	–1.7
	SA024	6.9	4.9
	SA025	–3.5	–4.7
	SA026	–3.5	–3.3
	SA027	2.5	1.4
	Mean	1.2	–1.0
	s.d.	3.7	3.4
$S_{60}N_{-60}$	SA021	–0.1	–4.0
	SA022	1.3	–3.3
	SA023	–6.7	–5.3
	SA024	2.0	–0.3
	SA025	–9.6	–9.7
	SA026	–9.1	–6.6
	SA027	1.4	–3.0
	Mean	–2.9	–4.6
	s.d.	5.2	2.9

7.2.3. Test conditions and processing strategies

In the two studies, listening effort was measured in bilateral listening for two spatial configurations of target and masker sources ($S_{15}N_{-15}$ and $S_{60}N_{-60}$), for four processing strategies in the Study 1 (STD, MOC1, MOC2, and MOC3) or two processing strategies in the Study 2 (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 72 effort estimates in total for Study 1 (2 spatial configurations × 4 strategies × 3 SNRs × 3 estimates per condition), and 36 estimates in total for Study 2 (2 spatial configurations × 2 strategies × 3 SNRs × 3 estimates per condition).

The target was always presented at an azimuth ipsilateral to the listener´s self-reported better ear, but the $S_{15}N_{-15}$ or $S_{60}N_{-60}$ nomenclature was chosen by convention (see Chapter 1).

7.2.4. 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 and effort estimate, conditions (spatial configurations and processing strategies) were administered in random order. Participants were given a brief break between test blocks.

7.2.5. Statistical analyses

For the two studies, we obtained the proportion of recognized words and the proportion of recalled words. An arcsine transformation[6] was applied to each of the two proportions to make them suitable for further statistical analyses (Studebaker, 1985; Studebaker et al., 1995):

$$T[\mathrm{AU}] = \arcsin\left(\sqrt{\frac{\mathrm{s}}{\mathrm{N+1}}}\right) + \arcsin\left(\sqrt{\frac{\mathrm{s+1}}{\mathrm{N+1}}}\right) \quad , \tag{7.1}$$

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

Because three measurements of recognized and recalled words were obtained per condition, a transformed proportion (in AU units) was calculated for each of the three measurements 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 $S_{60}N_{-60}$). Shapiro-Wilk tests were used to test if the distributions of recognized words, recalled words, and response times were normal (Gaussian). When this happened, parametric RMANOVAs were used to test for the effects of processing strategy, SNR, and their interaction on the transformed

---

[6] 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.

proportions of recognized and recalled words. When the distributions were not normal, 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).

7.3. RESULTS OF STUDY 1

7.3.1. Word recognition

Figure 7.1A illustrates the transformed proportion of recognized words (in AU units) in bilateral listening for the STD, MOC1, MOC2 and MOC3 processing strategies. Each data point is the mean AU score across six CI users (SA009, SA010, SA011, SA014, SA015 and SA016). Each panel is for a different spatial configuration: $S_{15}N_{-15}$ and $S_{60}N_{-60}$. The expected range of transformed proportions values would span 0.306 AU (corresponding to a proportion of 0) to 2.835 AU (corresponding to a proportion of 1).

One would expect word recognition scores to be highest in quiet than in noise at +5 dB SNR or at the individual SRT. Given that the individual SRTs were generally (very) negative across conditions (**Table 7.1**), one would also expect scores at +5 dB SNR to be similar to the scores in quiet or to be between the scores in quiet and scores at the individual SRT in noise. In addition, one would expect recognition scores in the SRT condition to be approximately constant across processing strategies because, by definition, the SRT condition implies equal (sentence) recognition across strategies. The overall pattern of results was consistent with these expectations.

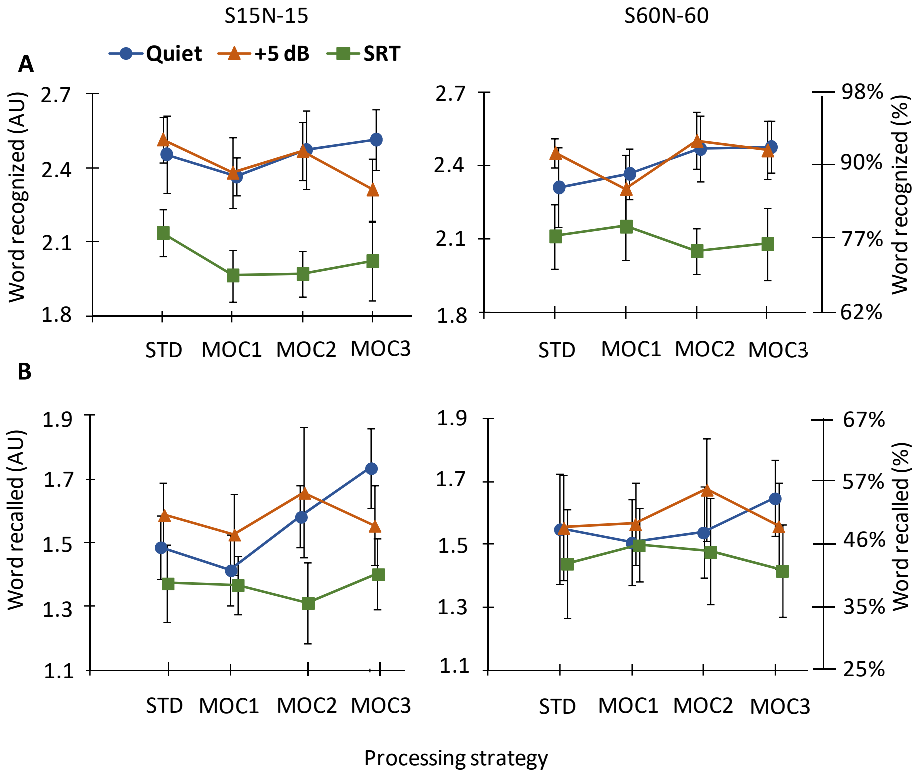

Figure 7.1. Mean proportions of recognized (**A**) and recalled words (**B**). The left and right panels illustrate results for the $S_{15}N_{-15}$ and $S_{60}N_{-60}$ spatial configurations, respectively. In each panel, the left axis shows the results expressed in AU units, and the right axis shows the corresponding scores in percentage. Each panel illustrates results for different sound-processing strategies (abscissae) and different SNRs, as indicated by the inset. Each data point is the mean score for six participants (and for three measurements per condition). Error bars indicate one standard error of the mean (N=6).

For the $S_{15}N_{-15}$ spatial configuration, Shapiro-Wilk tests showed that word recognition scores conformed to a normal distribution, thus it was justified to use a two-way RMANOVA to test for the effect processing strategy and SNR on word recognition scores. The RMANOVA revealed no significant effects of processing strategy [F(3,15)=2.070, p=0.147] or of the interaction between strategy and SNR [F(6,30)=0.733, p=0.627]. The effect of SNRs was statistically significant [F(2,10)=13.033, p=0.002]. Post-hoc pairwise comparisons with Bonferroni corrections revealed a significantly higher proportion of recognized words at +5 dB SNR than at the individual SRT (p=0.008). However, we found no statistically significant differences between the proportions of recognized words for other SNRs (quiet vs. +5 dB SNR, p=1.0; quiet vs. SRT, p=0.056). It was surprising that recognition scores in quiet were not significantly different (higher) from scores in the SRT condition, which was noisier, but the difference was close to being significant (p=0.056). This could be due to the small sample size and because participants had more practice in the SRT condition, which was passed last to participants (see Section 7.2.4).

For the $S_{60}N_{-60}$ spatial configuration, Shapiro-Wilk tests showed that word recognition scores for the MOC1 strategy at the SRT condition did not conform to a normal distribution (mean=2.151, s.d.=0.339, p=0.009), so it was justified to use a Friedman test to evaluate the effect of test condition on word recognition scores. The test revealed a statistically significant effects of test condition [$\chi^2(11)=27.849$, p=0.003] (note that there were 12 test conditions = 4 strategies × 3 SNRs). Post-hoc pairwise Wilcoxon signed-rank tests with Bonferroni corrections showed a higher proportion of recognized words with the MOC2 strategy in quiet than at the individual SRT (p=0.002), and at +5 dB SNR than at the individual SRT (p=0.002); and with the MOC3 strategy at +5 dB SNR than at the individual SRT (p=0.004). However, we found no statistically significant differences between any other pair of test conditions.

7.3.2. Word recall

Figure 7.1B shows the transformed proportion of recalled words (in AU units) in bilateral listening for the four processing strategies. Each data point is the mean score across the six CI participants.

For the $S_{15}N_{-15}$ spatial configuration, a two-way RMANOVA did not reveal an effect of processing strategy on word recall [F(3,15)=2.783, p=0.077]. It revealed a significant effect of SNR [F(2,10)=4.671, p=0.037]. A post-hoc analysis, using Bonferroni correction for multiple comparisons, revealed a significantly higher proportion of recalled words at +5 dB SNR than at the individual SRT (p=0.010). However, we found no statistically significant differences between the proportions of recalled words for other SNRs (quiet vs. +5 dB SNR, p=1.0; quiet vs. SRT, p=0.245). The RMANOVA also revealed a significant interaction between processing strategy and SNRs [F(6,30)=2.468, p=0.046]. Post-hoc pairwise comparisons with Bonferroni correction showed a significant effect of SNR for the MOC3 strategy and produced the following p values: p(quiet vs. +5 dB SNR) = 0.108; p(quiet vs. SRT) = 0.019; p(+5 dB SNR vs. SRT) = 0.060. In summary, this analysis showed that for the MOC3 strategy, word recall scores were better in quiet than in noise at the individual SRT and tended to be better at +5 dB SNR than in the SRT condition. In other words, this analysis showed that listening with MOC3 strategy required less effort in quiet than at +5 dB SNR or the SRT in noise. For the other processing strategies, the effect of SNR on word recall was not statistically significant.

For the $S_{60}N_{-60}$ spatial configuration, Shapiro-Wilk tests showed that word recalled scores for the STD strategy in quiet did not conform to a normal distribution (mean=1.549, s.d.=0.425, p=0.011), so it was justified to use a Friedman test to evaluate the effect of test condition on word recall scores. The test revealed no statistically significant differences in word recall across the 12 test conditions [$\chi^2(11)$=17.008, p=0.108].

7.4. RESULTS OF STUDY 2

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 time scores for the two strategies. Lastly, we report a correlation analysis between the two measures of listening effort (recall scores and response times).

7.4.1. Word recognition

Figure 7.2A illustrates the transformed proportion of recognized words (in AU units) in bilateral listening for the two processing strategies tested. Each data point is the mean AU score across seven bilateral CI users (recall that the score for each participant and test condition was the mean of three estimates). Each panel refers to different spatial configuration: $S_{15}N_{-15}$ and $S_{60}N_{-60}$.

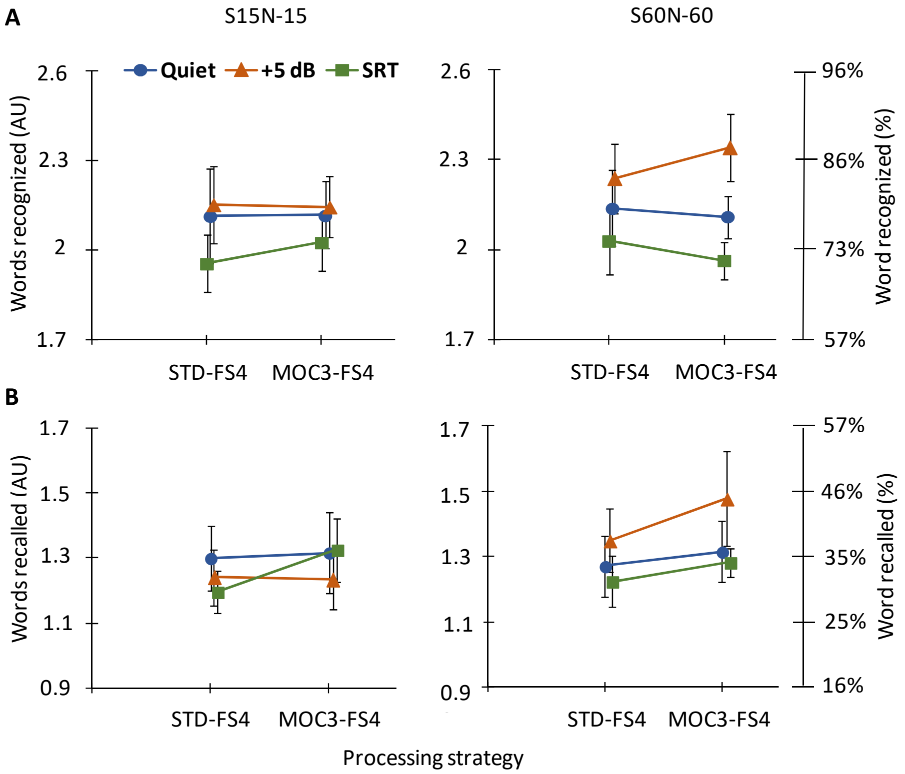

Figure 7.2. Mean proportion of recognized (**A**) and recalled words (**B**) for the STD-FS4 and MOC3-FS4 strategies, for the $S_{15}N_{-15}$ (**left**) and $S_{60}N_{-60}$ (**right**) spatial configurations. In each panel, the left axis shows the scores expressed in AU units, and the right axis shows corresponding scores in percentage. Each data point is the mean score for seven participants. Error bars indicate one standard error of the mean (N=7).

For the $S_{15}N_{-15}$ spatial configuration, a two-way RMANOVA revealed no significant effects of processing strategy [F(16)=0.110, p=0.752], SNR [F(2,12)=3.592, p=0.060], or interaction between strategy and SNR [F(2,12)=0.223, p = 0.804] 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 vs. +5 dB SNR, p=0.408; quiet vs. SRT, p=0.786).

7.4.2. Word recall

Figure 7.2B shows the transformed proportion of recalled words (in AU units) in bilateral listening for the STD-FS4 and MOC3-FS4 strategies. Each data point is the mean score across seven bilateral CI users (recall that the score for each participant and test condition was the mean of three estimates).

For the $S_{15}N_{-15}$ spatial configuration, A Friedman test revealed no statistically significant differences in word recall across the six test conditions [$\chi^2(5)=3.682$, p=0.596] (2 strategies × 3 SNRs).

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], or 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 and SNR conditions. In addition, there was no interaction between processing strategy and SNRs for either of the two spatial configurations tested.

7.4.3. Verbal response times

Figure 7.3 displays mean verbal response times as a function of processing strategy (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 times across seven BiCI users (note that the score for each participant and test condition was the mean of three estimates).

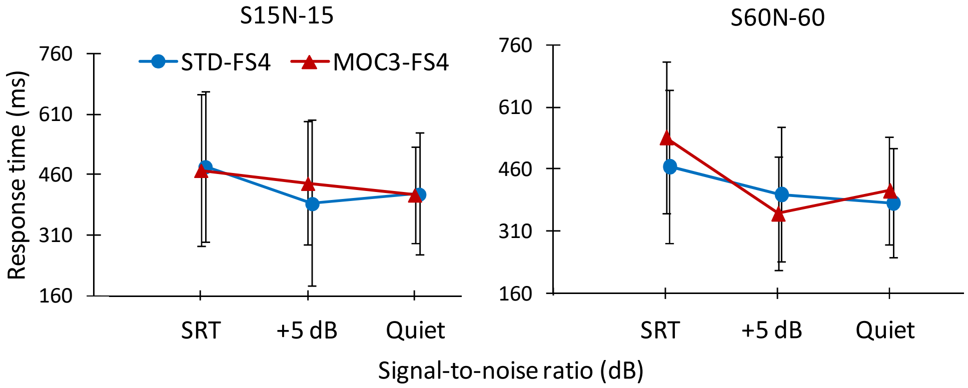

Figure 7.3. Mean verbal response time in the word recognition task averaged across subjects (N=7). The left and right panels illustrate response times for the $S_{15}N_{-15}$ and $S_{60}N_{-60}$ spatial configurations, respectively. Each panel illustrates response times for three different SNRs (abscissae) 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).

Response times tended to increase with decreasing the SNR (i.e., they tended to be shorter in quiet than for words in noise at the individual SRT). This is consistent with expectations because the individual SRTs were generally negative across conditions (**Table 7.2**) and made word recognition harder (**Fig. 7.2A**). In addition, response times were similar for the two processing strategies. Friedman tests revealed that response times were not statistically significantly different for any of 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.

7.4.4. Correlation between the two measures of listening effort

We conducted a correlation analysis to investigate if the two measures of listening effort used in Study 2 (word recall and verbal response time) reflected the same dimension of listening effort. **Figure 7.4** shows a plot of the number of recalled words (no arcsine transformation applied) against the verbal response time for each measurement. Note that the total number of data points in the figure (252) equals the product of 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, as one would expect. Spearman's rank correlation coefficient between the two variables was negative and statistically significant (ρ=–0.329, p<0.001). This suggests that the two measures partially reflected the same dimension and that both were sensitive to measure the effort. In other words, a greater number of recalled words and shorter response times probably reflected less listening effort.

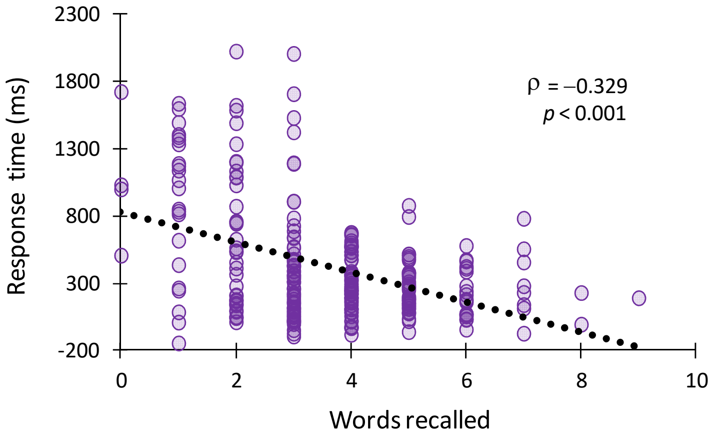

Figure 7.4. Correlation between the verbal response time and the number of recalled words. Data are pooled across seven participants, two processing strategies (STD-FS4 and MOC3-FS4), two spatial configurations ($S_{15}N_{-15}$ and $S_{60}N_{-60}$), three different SNRs (quiet, +5 dB SNR and individual 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.

7.5. DISCUSSION

We hypothesized that listening with the MOC strategies would require equal or less effort than listening with a strategy closer to the current clinical standard (the STD strategy). Study 1 was aimed at comparing the effort experienced by BiCI users when listening bilaterally to sounds processed through the STD, MOC1, MOC2, and MOC3 strategies, none of which involved FS4 processing. Listening effort was assessed using a dual-task paradigm, i.e., we measured word recognition (primary task) and word recall (secondary task). Study 2 was aimed at comparing the effort experienced by BiCI users when listening bilaterally with the MOC3-FS4 and STD-FS4 strategies. In Study 2, effort was assessed using the same dual-task paradigm as in Study 1 as well as using a single-task paradigm (verbal response time). In the two studies, effort was assessed for speech at three different SNRs (in quiet, at +5 dB SNR, and at the individual SNR for 50% sentence-in-noise recognition) and for two spatial configurations of the speech and noise sources ($S_{15}N_{-15}$ and $S_{60}N_{-60}$).

7.5.1. Assessment of listening effort with a dual-task paradigm

The hypothesis of the two studies was that recognizing speech in noise required the same or less effort with the MOC than with the STD strategy. Overall, the results of the two studies indicated that the proportion of recalled words was not significantly different for the two strategies. Assuming that the proportion of recalled words is a measure of effort, our results indicate that BiCI users experienced similar amounts of effort when listening with the MOC and the STD strategies. This is a positive finding, as participants were almost certainly more accustomed to listen to speech processed through the STD than the MOC strategies because the STD strategies were closer to those implemented in their clinical devices

In the two studies, word recognition and word recall scores tended to be better at +5 dB SNR than in quiet or at the individual SRT in noise, regardless of spatial configuration. This was probably because individual SRTs in noise were generally (very) negative for most participants (**Table 7.1 and 7.2**), and word recognition is harder at lower SNRs. The better scores a +5 dB SNR than in quiet could be due to participants being more experienced in the task at +5 dB SNR because this condition was administered after the quiet condition.

The dual task paradigm, as a method for assessing listening effort, has ecological validity since it assumes that people need to perform multiple tasks while listening in their daily lives (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 performance) are a direct measure of mental effort (McGarrigle et al., 2014).

In addition, it is well known that adults and children with hearing loss exert greater listening effort than listeners with normal hearing (Downs, 1982; Hicks and Tharpe, 2002; McCoy et al., 2005; Zekveld et al., 2011; Hornsby, 2013; Desjardins and Doherty, 2014). For this reason, it would be interesting to know to what extent the performance of the BiCI users compares with the performance of normal-hearing listeners in the same task. Such a comparison would inform of how far from 'normal' the performance of BiCI users is with and without the different strategies. As a first approximation, it would be interesting to test normal-hearing listeners using stimuli vocoded with the STD and MOC processing strategies.

7.5.2. Assessment of listening effort with verbal response times

In Study 2, we used the verbal response time as a complementary assessment of listening effort. For the two spatial configurations tested, we found no significant differences in response times across sound-processing strategies and SNRs. However, we saw that response times increased with decreasing SNR (**Fig. 7.3**), which seems reasonable because speech recognition gets harder with decreasing SNR. This is related to the assumption that increased task difficulty results in longer processing time, as more cognitive load is required to recognize and respond to stimuli. However, increased response times might not reflect greater effort. Increased task difficulty could result in increased effort to maintain the same level of performance, with no difference in response time despite increased effort (Bess and Hornsby, 2014).

The lack of a consensus on the most appropriate methodology to assess effort motivated us to use two different methodologies in Study 2. Overall, we found no significant differences in effort across the STD-FS4 and MOC3-FS4 strategies as assessed with either methodology. Still, the two measures reflected to some extent the expected changes in effort with corresponding changes in listening demand (e.g. in the more difficult or SRT-in-noise condition). This suggests that the two methods reflect a common index of listening effort.

7.5.3. Correlation between measures of effort

We found a correlation between the two measures of listening effort used in Study 2 (**Fig. 7.4**). We assumed that if the two methodologies were independent of the intelligibility performance, they would be reflecting in themselves the effort experienced by the listeners. However, word recognition scores were correlated with verbal response times (**Fig. 7.5A**) as well as with word recall scores (**Fig. 7.5B**) when the data were pooled across processing strategies, spatial configurations and SNRs. This could indicate that verbal response times and word recall scores are indirect measures of intelligibility, rather than of effort per se. This interpretation is supported by the fact that the study was double blind, that is, neither the experimenter nor the participants were biased to attend more strongly to sounds processed with any particular strategy, nor were they asked to respond as quickly as possible. In other words, participants probably performed the task devoting the same effort in all conditions.

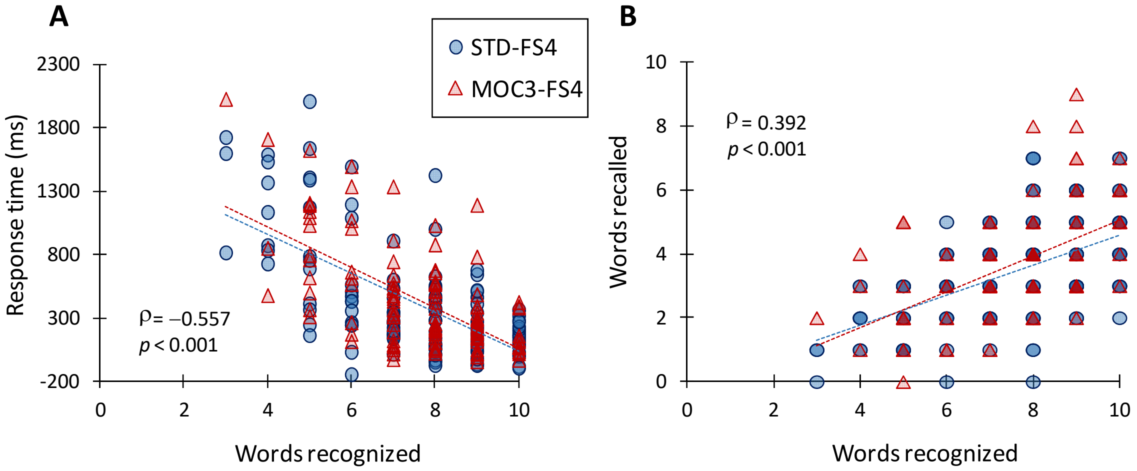

Figure 7.5. Correlation between verbal response time (**A**) and number of recalled words (**B**) with the number of recognized words. Data are pooled across the seven subjects, two processing strategies (STD-FS4 and MOC3-FS4), two different spatial configurations ($S_{15}N_{-15}$ and $S_{60}N_{-60}$), three different SNRs (quiet, +5 dB SNR and individual SRT in noise), and three measurements per condition. Also shown are Spearman rank correlation coefficients (ρ) and the probability of observing those values by chance (p).

7.5.4. Limitations

The principal limitation of the two studies was the lack of a standardized test procedure to measure listening effort. We used the behavioral dual-task approach to assess effort because it provides a form of ecological validity to the experimental procedure in that dual tasking is often required when processing speech in real-life situations (Gagné et al., 2017). It would be interesting to assess effort using objective (physiological) methods (e.g. pupillometry) in future studies.

In the two studies, the sample size was small and heterogeneous in age, duration of implant use, and deafness etiology (see **Tables 3.1 and 3.2**). These factors may have increased the variability in performance and contributed to the lack of significant effects of the effort experienced by the participants with each strategy. The type of background noise influences listening effort (Larsby et al., 2005). The stationary noise used in the present studies is commonly used in research, but it would be interesting to evaluate the effort experienced by CI users with the MOC strategy using more natural noises (e.g. babble noise). The very limited experience of participants with the MOC strategies (they used them only for laboratory tests and for four days) may have also increased the variability in the data.

7.6. CONCLUSIONS

The two studies reported in this chapter were a first step toward investigating whether listening with various MOC strategies is more or less effortful than listening with more conventional sound-processing strategies. We found that:

(1) In quiet, word recall scores were better with the MOC3 than with the STD, MOC1 or MOC2 strategies. At +5 dB SNR or at the individual SRT in noise, word recall scores were not statistically different for any pair of strategies.

(2) Word recall scores and verbal response times were not significantly different for the STD-FS4 and MOC3-FS4 processing strategies.

(3) There was a correlation between the number of recalled words and the verbal response time. This suggests that those two measures possibly reflected changes in listening effort corresponding to changes in listening demands.

(4) Altogether, the data suggested that listening with the MOC strategies requires less than or comparable effort as listening with more conventional sound-processing strategies.

8.

GENERAL DISCUSSION

The MOC strategy is a binaural CI sound-processing strategy with dynamic, contralateral control of acoustic-to-electric compression inspired by the natural MOCR (Lopez-Poveda et al., 2016a, 2016b, 2017; Lopez-Poveda and Eustaquio-Martín, 2018). The aim of the present work was to experimentally evaluate various different implementations of the MOC strategy designed to reflect more realistically the inhibitory characteristics of the natural MOCR. The evaluation consisted of comparing the performance of CI users on several hearing tests with the MOC strategy against that with a STD strategy that involved using two independently functioning devices with fixed compression. Evaluation tests included speech-in-noise recognition for multiple spatial configurations of the target and masker sources, sound-source localization in the horizontal plane, and listening effort during speech recognition. The tested MOC and STD strategies included versions intended to disregard as well as to preserve TFS speech cues.

In **Chapter 4**, it has been shown that the recognition of sentences in noise 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. This strategy 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 strategy tested elsewhere (Lopez-Poveda et al., 2016a, 2017).

In **Chapter 5**, it has been shown that, compared to the STD strategy, the MOC1 strategy (with fast control of compression and greater inhibition in the higher frequency channels), improves the localization of brief (200 ms) wideband (125-6000 Hz) noise bursts in a virtual horizontal plane, because it enhances the head-shadow ILDs. More realistic implementations of the MOC strategy with slower control of compression, and/or slightly greater inhibition in the lower frequency channels (MOC2 and MOC3 strategies) also provide larger ILDs than the STD strategy but for sufficiently long stimuli (>300 ms).

In **Chapter 6**, it has been shown that the benefits of MOC3 processing for the recognition of sentences in noise are preserved when the sound-coding involves FS4 processing intended to preserve some speech TFS cues. More specifically, SRTs for sentences in noise were equal or slightly better with the MOC3-FS4 than with the STD-FS4 strategy in unilateral and bilateral listening modes across a wider range of speech levels (–28, –38 and –48 dB FS) and maskers types (SSN and iFFM).

Lastly, in **Chapter 7**, it has been shown that recognizing speech in noise is as effortful with the MOC strategies as it is with the more conventional STD strategies.

8.1. GREATER OVERALL BENEFITS WITH MORE REALISTIC IMPLEMENTATIONS OF THE MOC STRATEGY

One of the main findings of this work is that the benefits of the originally proposed MOC1 strategy (Lopez-Poveda et al., 2016a, 2017) can be enhanced and its shortcomings overcome by using more realistic implementations of MOC processing (the MOC2 and MOC3 strategies) (**Chapter 4**). Overall, speech-in-noise recognition was better 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.

It has been shown that compared to the STD strategy, the MOC1 strategy enhances ILD cues (**Figs. 5.4** and **5.5**) and improves sound source localization for short (200 ms) acoustic stimuli (**Figs. 5.7** and **5.8**). More realistic implementations of the MOC strategy with slower (longer) time constant of integration (MOC2) and with greater inhibition in the lower than in the higher frequency channels (MOC3) can also enhance ILD cues for stimuli that are sufficiently long to fully activate and deactivate the contralateral inhibition in these two strategies (**Fig. 5.3**).

8.2. LISTENING EFFORT IS SIMILAR WITH THE MOC AND STD STRATEGIES

We have shown that both speech-in-noise recognition and sound source localization are overall equal or better with the MOC strategies than with strategies closer to the current clinical standard. We have also shown that recognizing speech (in quiet or in noise) with the MOC strategies requires as much effort as listening with the STD strategies. Together, this suggests that MOC processing can improve hearing without increasing listening effort.

In this thesis, listening effort has been assessed using word recall as well as the verbal response time in a word recognition task. A motivation for using two different methods was to investigate if the two metrics reflect a common dimension of listening effort. It has been shown (**Chapter 7**) that the two metrics correlate with each other, and scores 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 of effort and both are sensitive to effort. However, it is important to note that a variety of methods is currently used in research settings to assess listening effort, including behavioral and physiological measures (see the review in **Section 3.6**). It is not clear if and to what extent the different methodologies tap into the same construct of effort because they rarely correlate with each other (Alhanbali et al., 2019). Therefore, the absence of a standard methodology for assessing listening effort limits the capacity to confirm that the different measures are related to the same concept of listening effort (McGarrigle et al., 2014; Alhanbali et al., 2019).

8.3. COMPARISON WITH OTHER BINAURAL ALGORITHMS OR PROCESSING STRATEGIES

There exist other sound-processing approaches aimed at bringing the hearing performance of BiCI users closer to that of listeners with normal hearing. Because the use of independent sound processors with independent compression at the two ears can distort ILD cues and degrade speech-in-noise intelligibility (e.g. Wiggins and Seeber, 2013), one approach consists of using linked (equal) AGC across the ears (e.g., Potts et al., 2019; Spencer et al., 2018). Compared to 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 pre-processing 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 pre-processing 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).

On the other hand, several ILD-enhancement methods have been recently proposed for CIs. Moore et al. (2016) proposed to enhance ILDs at low frequencies (≤1500 Hz; the frequency range where HRTF ILDs are smaller and ITDs are greater) by 'mapping' inter-aural phase differences into ILDs. The method was expected to create a (correct) perception of sound source location, even if the listener is insensitive to the corresponding ITD, as might be the case for typical BiCI users. The method was tested on bilateral hearing-aid users using simulated hearing aids. The algorithm did not improve the localization of noise sources but improved the localization of speech sources by a few degrees at some azimuths (the mean improvement across azimuths was not reported). Dieudonné and Francart (2018) proposed to enhance head-shadow ILDs using a fixed beamformer with contralateral attenuation in each ear. The method was tested on normal hearing listeners simulating bimodal stimulation (i.e., listening with a simulated CI in one ear and a simulated hearing loss in the other ear). Root-mean-square localization angle errors improved from 50.5° without the beamformer to 26.8° with the beamformer. While potentially useful for BiCI users, as far as is known, neither method has yet been tested on BiCI users. Therefore, it is not possible to directly compare their benefit with the benefit provided by MOC processing.

Other ILD-enhancement approaches have been specifically designed and tested for BiCI users. For example, Francart et al. (2011) proposed an ILD enhancement algorithm for bimodal CI users that improved individual lateralization scores in a virtual sound field by 4° to 10°. The mean angle error decreased from 28.4° without ILD enhancement to 20.6° with enhancement. Brown (2018) proposed a sound-processing strategy intended to provide BiCI with larger than normal ILD cues. Mean angle error improved from 31.0° without enhancement to 12.8° with ILD enhancement. These ILD-enhancement strategies provide better absolute lateralization scores and larger improvements with respect to the reference condition than the MOC1 strategy (mean angle error = 22.7°, improvement re STD = 2.5°; **Fig 5.7**). We note, however, that the enhancement of ILD cues in the MOC strategies is an emergent property of MOC processing rather than an intended effect (Lopez-Poveda, 2015).

Overall, a direct comparison of the benefit provided by the approaches just described 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 SNR across the spatial configurations tested in **Chapter 4**) appears smaller than the benefit provided by some of those approaches. Binaural pre-processing 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 pre-processing, and probably less exchange of data between the ears. These characteristics make the MOC strategy suitable for implementation in clinical devices.

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 spatial configuration (**Fig. 4.7**). This possibly reflects envelope enhancement due to the use of an overall more linear acoustic-to-electric maplaw, and/or neural 'antimasking' associated to a reduced stimulation. Other benefits of MOC processing, 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) AGC and natural ILDs may be somewhat restored by using linked AGC (Wiggins and Seeber, 2013), MOC processing might provide larger benefits when used in combination with linked AGC. On the other hand, MOC processing can be theoretically implemented with *any* CI sound-coding strategy that does not already utilize dynamic back-end compression. Indeed, the study reported in **Chapter 6** demonstrates that MOC3 processing can be combined with FS4 processing and that when used in combination with linked AGC, it produces equal or better SRTs in noise than a STD strategy. The study in question, 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 with other sound-coding strategies.

8.4. LIMITATIONS

Given the limited number of sentence lists in the HINT and Sharvard corpora used here to evaluate the MOC strategies, sentence lists had to be used multiple times to complete a protocol. It is likely that participants learnt some of the sentences during testing. As a result, the reported SRTs are probably lower than they would have been if the speech material had not been used repeatedly. We are confident, however, that re-using the sentences did not contribute to the reported differences in SRTs across strategies (or spatial configurations) because anyone 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.

At the time when the tests were conducted, all participants had a long, daily experience with an audio-coding strategy similar to the STD or STD-FS4. By contrast, their experience with the MOC strategies was limited to a few days during the test sessions. For this reason, the found benefits of MOC processing for speech-in-noise recognition or sound source localization are striking. Furthermore, the very limited experience of the participants with the use of MOC strategies could have increased the variability in scores with these strategies. For CI users, speech recognition can improve significantly over time and with training (e.g., Dorman and Spahr, 2006) and some benefits of bilateral implantation, such as squelch, are seen only one year after the start of CI use (e.g., Buss et al., 2008). For this reason, it is conceivable that the benefits from the MOC strategy could become larger with practice and/or a sustained use of the strategy.

On the other hand, in the localization study the stimulus duration (200 ms) was shorter than the time required for a full activation (and deactivation) of contralateral inhibition in the MOC2 or MOC3 strategies (**Fig. 5.3**). As a result, the ILDs with those strategies and stimuli were probably smaller than they would have been for longer stimuli. Therefore, it is also possible that the use of longer stimuli could improve localization performance with the more realistic implementations of the MOC 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 measure 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 can be affected by individual differences in aspects such as task engagement and motivation (Alhanbali et al., 2019). Studies using dual-task paradigms have demonstrated that behavioral measures suffer from imprecision and are difficult to compare results across studies (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 identify whether participants use all their cognitive capacity or not. Further, it is not possible to know with certainly if the participant always prioritizes the performance of the primary task (Alhanbali et al., 2019).

8.5. WHICH MOC STRATEGY?

Altogether, the results of the different experimental evaluations described in this thesis show that the binaural MOC strategy can facilitate the localization of sound sources and the recognition of in noise without increasing listening effort. However, we have observed that some MOC strategies provide a greater benefit compared to others depending on the listening task, the duration of the stimulus, and the type of stimulus (mainly the type of masker) used.

In addition to improving sound source lateralization scores, the MOC1 strategy can also improve the intelligibility of speech when the target source is presented in competition with another talker (Lopez-Poveda et al., 2017) or with a source of steady-state noise (Lopez-Poveda et al., 2016a). As explained earlier, however, the MOC1 strategy has potential drawbacks: (1) it can reduce the speech information in the ear opposite to the target source (i.e., the ear with the worse acoustic SNR), which could potentially hinder intelligibility in unilateral listening when the implant ear has the worse acoustic signal-to-noise ratio; and (2) the mutual inhibition between the pair of processors can decrease the overall stimulation levels and thus audibility, which could hinder intelligibility in bilateral or unilateral listening when the two CIs (or processors) have input signals with identical levels. (The latter drawback is less of a concern in realistic listening conditions because any asymmetrical placement of the CI microphones would suffice for the levels of the input signals to be different). We have shown that the MOC2 and/or the MOC3 strategy could overcome the two drawbacks (Chapter 4) but their control of compression is too slow to enhance ILDs for brief sounds (< 1 second in duration) such as those employed in the present localization tests (Chapter 5). As discussed in Chapter 5, however, the MOC2 strategy and to some extent also the MOC3 strategy could provide ILDs closer to those of the MOC1 strategy for longer stimuli.

The more realistic MOC3 strategy solved the shortcomings of the original MOC1 strategy and provided overall better speech-in-noise recognition. In addition, the MOC2 and MOC3 strategies produced a statistically significant binaural advantage, which did not occur with the STD or MOC1 strategy (Chapter 4). This could be because the more realistic implementations of the MOC strategy (MOC2 and MOC3) slightly improved the speech information in the ear with the worse acoustic SNR and/or probably transmitted more natural binaural information [see **Figure 4.3** and Lopez-Poveda and Eustaquio-Martin (2018)].

Altogether, it seems that the faster contralateral inhibition (i.e., as in the MOC1 strategy described by Lopez-Poveda et al., 2016a, 2017) might be more advantageous for speech in competition with fluctuating maskers and localization of short and long stimuli while slower contralateral inhibition (as in MOC2 or MOC3 strategies) might be more advantageous for speech presented in competition with steady-state maskers and localization of longer stimuli.

In summary, all MOC strategies hold potential for improving some aspects of the hearing of BiCI users. However, it will be important to continue this research to elucidate which implementation and parameters provide the greatest overall benefit to CI users.

8.6. OUTLOOK

The average benefits of MOC processing hold for many individual CI users. 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 the tests. As discussed earlier, the potential benefits from MOC processing could become larger with practice and/or a sustained use of the MOC strategies. To quantify the actual benefits of the MOC strategy, it would be important and useful to provide participants with the MOC strategy in portable hardware format, so that they can use it in their daily life and get used to it. This hardware would also allow carrying out tests like the ones reported here but in 'true' rather than simulated free-field listening conditions.

It would also be important to test the MOC strategy in more realistic listening environments. Because everyday hearing is dynamic (i.e., people and objects are mobile) and MOC processing is also dynamic, it would be interesting to evaluate hearing performance with the MOC strategies using moving sound sources in real-world listening scenarios.

On the other hand, in real life, listeners often require following two concurrent competing signals (e.g., two simultaneous talkers). Given that the MOC strategy enhances the ILDs and can potentially improve the spatial segregation of spatially separated sound sources (see **Fig. 4.3** and Fig. 3 in Lopez-Poveda et al., 2016b), it would also be interesting to assess to what extent MOC processing can facilitate following two simultaneous conversations in realistic environments.

Because the MOC2 and MOC3 strategies theoretically enhances ILDs for sufficiently long stimuli, it would also be interesting to extend the sound localization study reported in **Chapter 5** to stimuli that are sufficiently long (>300 ms) to fully activate and deactivate contralateral inhibition. This could help elucidating which implementation and parameters provide the greatest overall benefit to the patient.

Additionally, it would be interesting to further improve the parameters and implementations of the MOC strategy. For example, to investigate whether contralateral inhibition with intermediate time constants improves speech-in-noise intelligibility for all types of maskers.

Lastly, it would also be worthwhile evaluating listening effort during a speech-in-noise recognition with the MOC strategies using physiological rather than behavioral methods, such as, for example, pupillometry.

CONCLUSIONS

The four studies reported in this thesis were aimed at experimentally evaluating various alternative implementations of the MOC strategy designed to reflect more or less realistically the time course and magnitude of contralateral MOCR inhibition. The main conclusions are:

(1) Compared to using two independently functioning sound processors (a strategy similar to the current clinical STD), a MOC strategy with fast control of compression and greater inhibition in the higher-frequency than in the lower-frequency channels (MOC1), slightly improves the localization of shorter (200 ms) stimuli in a virtual horizontal plane. However, MOC implementations that involve slower control of compression, and/or slightly greater inhibition in the lower-frequency than in the higher-frequency channels (MOC2 and MOC3 strategies) also provide theoretical benefits for sufficiently long stimuli (>1 s).

(2) Speech-in-noise recognition is overall better with the more realistic MOC3 strategy. This strategy maintains the benefits of the originally proposed MOC1 strategy over the STD strategy for spatially separated speech and noise sources and extend those benefits to additional spatial configurations. In addition, the MOC3 strategy provides a significant binaural advantage, which is not the case for the STD or the MOC1 strategy.

(3) The MOC3 strategy combined with FS4 processing (the MOC3-FS4 strategy) produces equal or better speech-in-noise recognition than the STD-FS4 strategy in unilateral and bilateral listening for a reasonably wide range of speech levels (–28, –38 and –48 dB FS), for multiple spatial configurations, and for steady-state and fluctuating maskers.

(4) Bilateral CI users experience approximately the same listening effort during a word-in-noise recognition task for sounds processed with the various STD and MOC strategies tested here.

(5) Altogether, the present studies show that MOC processing can improve the localization of sound sources in quiet and the recognition of speech in noise without increasing listening effort.

REFERENCES

Adiloğlu K, Kayser H, Baumgärtel RM, Rennebeck S, Dietz M, Hohmann V. (2015). A binaural steering beamformer system for enhancing a moving speech source. *Trends in Hearing* 19:1-13.

Aguilar E, Eustaquio-Martin A, Lopez-Poveda EA. (2013). Contralateral efferent reflex effects on threshold and suprathreshold psychoacoustical tuning curves at low and high frequencies. *Journal of the Association for Research in Otolaryngology* 14:341-357.

Aguilar E, Johannesen PT, Lopez-Poveda EA. (2015). Contralateral efferent suppression of human hearing sensitivity. *Frontiers in Systems Neuroscience* 8:1-8.

Akeroyd MA. (2006). The psychoacoustics of binaural hearing. *International Journal of Audiology* 45:25-33.

Alhanbali S, Dawes P, Lloyd S, Munro KJ. (2017). Self-reported listening-related effort and fatigue in hearing-impaired adults. *Ear and Hearing* 38:e39-e48.

Alhanbali S, Dawes P, Millman RE, Munro KJ. (2019). Measures of listening effort are multidimensional. *Ear and Hearing* 40:1084-1097.

Aquino-Acevedo T. (2016). Caracterización del test matricial de Oldenburg en Castellano para la valoración de la discriminación verbal en ruido en usuarios de implante coclear. Master Thesis. Máster en Trastornos de la Comunicación. University of Salamanca, Salamanca, Spain.

Arlinger S. (2003). Negative consequences of uncorrected hearing loss—a review. *International Journal of Audiology* 42:17-20.

Arsenault MD, Punch JL. (1999). Nonsense-syllable recognition in noise using monaural and binaural listening strategies. *Journal of Acoustical Society of America* 105:1821-1830.

Aubanel V, García Lecumberri ML, Cooke M. (2014). The Sharvard corpus: A phonemically-balanced Spanish sentence resource for audiology. *International Journal of Audiology* 53:633-638.

Avan P, Giraudet F, Büki B. (2015). Importance of binaural hearing. *Audiology and Neurotology* 20(suppl 1):3–6.

Backus BC, Guinan JJ (2006). Time-course of the human medial olivocochlear reflex. *Journal of Acoustical Society of America* 119:2889-2904.

Baskent D, Gaudrain E. (2016). Perception and Psychoacoustics of Speech in Cochlear Implant Users. *Scientific Foundations of Audiology. Perspectives from Physics, Biology, Modelling, and Medicine*, 185-320.

Baumgärtel RM, Hu H, Krawczyk-Becker M, Marquardt D, Herzke T, Coleman G, (…) and Dietz M. (2015a). Comparing binaural pre-processing strategies II: Speech intelligibility of bilateral cochlear implant users. *Trends in Hearing* 19:1-18.

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

Bernarding C, Strauss DJ, Hannemann R, Seidler H, Corona-Strauss FI. (2013). Neural correlates of listening effort related factors: Influence of age and hearing impairment. *Brain Research*

Bulletin 91:21–30.

Berrettini S, Baggiani A, Bruschini L, Cassandro E, Cuda D, Filipo R, (…) Forli F. (2011). Systematic review of the literature on the clinical effectiveness of the cochlear implant procedure in adult patients. *Acta Otorhinolaryngologica Italica : Organo Ufficiale Della Societa Italiana Di Otorinolaringologia e Chirurgia Cervico-Facciale* 31:299-310.

Bess FH, Hornsby BW. (2014). Commentary: Listening can be exhausting—Fatigue in children and adults with hearing loss. *Ear and hearing* 35:592.

Blauert J. (1997). Spatial hearing: The Psychophysics of Human Sound Localization, second ed. MIT Press, Cambridge.

Bond M, Mealing S, Anderson R, Elston J, Weiner G, Taylor RS, (…) Stein K. (2009). The effectiveness and cost-effectiveness of cochlear implants for severe to profound deafness in children and adults: A systematic review and economic model. *Health Technology Assessment* 13:1-330.

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 and Hearing* 27:608-618.

Boyle PJ, Büchner A, Stone MA, Lenarz T, Moore BC. (2009). Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants. *International journal of audiology* 48:211-221.

Bronkhorst AW, Plomp R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *Journal of Acoustical Society of America* 83:1508-1516.

Brown CA. (2018). Corrective binaural processing for bilateral cochlear implant patients. *PLoS ONE* 13:1–18.

Brown GJ, Ferry RT, Meddis R. (2010). A computer model of auditory efferent suppression: Implications for the recognition of speech in noise. *Journal of the Acoustical Society of America* 127:943–954.

Brown KD, Balkany TJ. (2007). Benefits of bilateral cochlear implantation: A review. *Current Opinion in Otolaryngology and Head and Neck Surgery* 15:315-318.

Brown MC, Nuttall AL, Masta RI. (1983). Intracellular recordings from cochlear inner hair cells: effects of stimulation of the cross olivocochlear efferents. *Science* 222:69-72.

Brown MC, Nuttall AL. (1984). Efferent control of cochlear inner hair cell responses in the guinea pig. *Journal of Physiology* 354:625-646.

Brownell WE. (1990). Outer hair cell electromotility and otoacoustic emissions. *Ear and Hearing* 11:82-92.

Buechner A, Beynon A, Szyfter W, Niemczyk K, Hoppe U, Hey M, (...) and Zarowski A. (2011). Clinical evaluation of cochlear implant sound coding taking into account conjectural masking functions, MP3000™. *Cochlear Implants International* 12:194-204.

Buss E, Pillsbury HC, Buchman CA, Pillsbury CH, Clark MS, Haynes DS, (…) Barco AL. (2008). Multicenter U.S. bilateral MED-EL cochlear implantation study: Speech perception over the first year of use. *Ear and Hearing* 29:20-32.

Cárdenas MR, Marrero V (1994). *Cuaderno de logoaudiometría. Madrid: Universidad Nacional de Educación a Distancia (UNED).*

Ching TYC, Van Wanrooy E, Dillon H. (2007). Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review. *Trends in Amplification* 11:161-192.

Cooper NP, Guinan JJ. (2003). Separate mechanical processes underlie fast and slow effects of medial olivocochlear efferent activity. *Journal of Physiology* 548:307-312.

Cooper NP, Guinan JJ. (2006). Efferent-mediated control of basilar membrane motion. *Journal of Physiology* 576:49–54.

Crathorne L, Bond M, Cooper C, Elston J, Weiner G, Taylor R, Stein K. (2012). A systematic review of the effectiveness and cost-effectiveness of bilateral multichannel cochlear implants in adults with severe-to-profound hearing loss. *Clinical Otolaryngology* 37:342-354.

Dallos P, Fakler B. (2002). Prestin, a new type of motor protein. *Nature Reviews Molecular Cell Biology* 3:104-111.

Dallos P, Wu X, Cheatham MA, Gao J, Zheng J, Anderson CT, (...) and He DZ. (2008). Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. *Neuron* 58:333-339.

Dallos P, Zheng J, Cheatham MA. (2006). Prestin and the cochlear amplifier. *Journal of Physiology* 576:37-42.

Dallos P. (2008). Cochlear amplification, outer hair cells and prestin. *Current Opinion in Neurobiology* 18:370-376.

Desjardins JL, Doherty KA. (2013). Age-related changes in listening effort for various types of masker noises. *Ear and Hearing* 34:261-272.

Desjardins JL. (2015). The effects of hearing aid directional microphone and noise reduction processing on listening effort in older adults with hearing loss. *Journal of the American Academy of Audiology* 27:29-41.

Di Lella F, Bacciu A, Pasanisi E, Vincenti V, Guida M, Bacciu S. (2010). Main peak interleaved sampling (MPIS) strategy: effect of stimulation rate variations on speech perception in adult cochlear implant recipients using the Digisonic SP cochlear implant. *Acta Oto-laryngologica* 130:102-107.

Dieudonné B, Francart T. (2018). Head shadow enhancement with low-frequency beamforming improves sound localization and speech perception for simulated bimodal listeners. *Hearing Research* 363:78–84.

DiGiovanni JJ, Riffle TL, Nagaraj NK. (2017). Attention-controlled working memory measures to assess listening effort. *Clinical Archives of Communication Disorders* 2:163-177.

Dorman MF, Loiselle L, Stohl J, Yost WA, Spahr A, Brown C, Cook S. (2014). Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear and Hearing* 35:633-640.

Dorman MF, Loiselle LH, Cook SJ, Yost WA, Gifford RH. (2016). Sound source localization by normal-hearing listeners, hearing-impaired listeners and cochlear implant listeners. *Audiology and Neurotology* 21:127-131.

Dorman MF, Spahr AJ. (2006). Speech perception by adults with multichannel cochlear implants. In: SB Waltzman, JT Roland (eds.), *Cochlear Implants*, 2nd ed., (Thieme, New York), pp. 193-204.

Downs DW. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders* 47:189-193.

Dubno JR, Ahlstrom JB, Horwitz AR. (2008). Binaural advantage for younger and older adults with normal hearing. *Journal of Speech, Language, and Hearing Research* 51:539-556.

Eapen RJ, Buss E, Adunka MC, Pillsbury III HC, Buchman CA. (2009). Hearing-in-noise benefits after bilateral simultaneous cochlear implantation continue to improve 4 years after implantation. *Otology and Neurotology* 30:153-159.

Eckert MA, Teubner-Rhodes S, Vaden KI. (2016). Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear and Hearing* 37:101S-110S.

Eddins AC, Ozmeral EJ, Eddins DA. (2018). How aging impacts the encoding of binaural cues and the perception of auditory space. *Hearing Research* 369:79-89.

Francart T, Lenssen A, Wouters J. (2011). Enhancement of interaural level differences improves sound localization in bimodal hearing. *Journal of the Acoustical Society of America* 130:2817-2826.

Francart T, Lenssen A, Wouters J. (2014). Modulation enhancement in the electrical signal improves perception of interaural time differences with bimodal stimulation. *Journal of the Association for Research in Otolaryngology* 15:633-647.

Fraser S, Gagné J-P, Alepins M, Dubois P. (2009). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. *Journal of Speech, Language, and Hearing Research* 53:18-33.

Fu Q-J, Shannon RV. (1998). Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners. *Journal of Acoustical Society of America* 104:2570-2577.

Gagné JP, Besser J, Lemke U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: A review. *Trends in Hearing* 21:1–25.

Gantz BJ, Tyler RS, Rubinstein JT, Wolaver A, Lowder M, Abbas P, (…) Preece JP. (2002). Binaural cochlear implants placed during the same operation. *Otology and Neurotology* 23:169-180.

Gardner WG, Martin KD. (1995). HRTF measurements of a KEMAR. *Journal of the Acoustical Society of America* 97:3907-3908.

Gatehouse S, Gordon J. (1990). Response times to speech stimuli as measures of benefit from amplification. *British Journal of Audiology* 24:63-68.

Gatehouse S, Noble W. (2004). The Speech, spatial and qualities of hearing scale (SSQ). *International Journal of Audiology* 43:85-99.

Goodman A. (1965). Reference zero levels for pure-tone audiometer. *American Speech-Language-Hearing Association* 72:347-350.

Gosselin PA, Gagné JP. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech, Language, and Hearing Research* 54:944-958.

Gosselin PA, Gagné JP. (2010). Use of a dual-task paradigm to measure listening effort. *Revue Canadienne d'Orthophonie et d'Audiologie* 34:43-51.

Goupell M J, Stakhovskaya OA, Bernstein JG. (2018). Contralateral interference caused by binaurally presented competing speech in adult bilateral cochlear-implant users. *Ear and Hearing* 39:110-123.

Goupell MJ, Litovsky RY. (2015). Sensitivity to interaural envelope correlation changes in bilateral cochlear-implant users. *Journal of the Acoustical Society of America* 137:335-349.

Goupell MJ, Stoelb C, Kan A, Litovsky RY. (2013). Effect of mismatched place-of-stimulation on the salience of binaural cues in conditions that simulate bilateral cochlear-implant listening. *Journal of the Acoustical Society of America* 133:2272-2287.

Goupell MJ, Stoelb CA, Kan A, Litovsky RY. (2018). The effect of simulated interaural frequency mismatch on speech understanding and spatial release from masking. *Ear and Hearing* 39:895-905.

Goupell MJ. (2015). Interaural correlation-change discrimination in bilateral cochlear-implant users: Effects of interaural frequency mismatch, centering, and age of onset of deafness. *Journal of the Acoustical Society of America* 137:1282-1297

Grantham DW, Ashmead DH, Ricketts TA, Labadie RF, Haynes DS. (2007). Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. *Ear and Hearing* 28:524-541.

Guinan JJ. (2006). Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. *Ear and Hearing* 27:589-607.

Gustafson SJ, Key AP, Hornsby BW, Bess FH. (2018). Fatigue related to speech processing in children with hearing loss: Behavioral, subjective, and electrophysiological measures. *Journal of Speech, Language, and Hearing Research* 61:1000-1011.

Hawley ML, Litovsky RY, Culling JF. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *Journal of the Acoustical Society of America* 115:833-843.

Hicks CB, Tharpe AM. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research* 45:573-584.

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 in Amplification* 10:201-219.

Hochmuth S, Brand T, Zokoll MA, Castro FZ, Wardenaga N, Kollmeir B. (2012). A Spanish matrix sentence test for assessing speech reception thresholds in noise. *International Journal of Audiology* 51:536-544.

Holube I, Fredelake S, Vlaming M, Kollmeier B. (2010). Development and analysis of an international speech test signal (ISTS). *International Journal of Audiology* 49:891-903.

Hornsby BW. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing* 34:523–534.

Houben R, Van Doorn-Bierman M, Dreschler WA. (2013). Using response time to speech as a measure for listening effort. *International Journal of Audiology* 52:753-761.

Howard CS, Munro KJ, Plack CJ. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology* 49:928-932.

Hua H, Emilsson M, Ellis R, Widén S, Möller C, Lyxell B. (2014). Cognitive skills and the effect of noise on perceived effort in employees with aided hearing impairment and normal hearing. *Noise and Health* 16:79-88.

Hua H, Karlsson J, Widén S, Möller C, Lyxell B. (2013). Quality of life, effort and disturbance perceived in noise: A comparison between employees with aided hearing impairment and normal hearing. *International Journal of Audiology* 52:642-649.

Huarte A. (2008). The Castilian Spanish hearing in noise test. *International Journal of Audiology* 47:369-370.

Hughes KC, Galvin KL. (2013). Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing. *Cochlear Implants International* 14:121-129.

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. *American Journal of Audiology* 24:419-431.

Jones H, Kan A, Litovsky RY. (2014). Comparing sound localization deficits in bilateral cochlear-implant users and vocoder simulations with normal-hearing listeners. *Trends in Hearing* 18:1-16.

Kahneman D. (1973). *Attention and Effort.* Englewood Cliffs, New Jersey.

Kan A and Litovsky RY. (2015). Binaural hearing with electrical stimulation. *Hearing Research* 322:127-137.

Kan A, Jones HG, Litovsky RY. (2015). Effect of multi-electrode configuration on sensitivity to interaural timing differences in bilateral cochlear-implant users. *Journal of the Acoustical Society of America* 138:3826-3833.

Kan A. (2018). Improving speech recognition in bilateral cochlear implant users by listening with the better ear. *Trends in Hearing* 22:1-11.

Kawase T, Ogura M, Sato T, Kobayashi, T, Suzuki Y. (2003). Effects of contralateral masking on threshold. *Tohoku Journal of Experimental Medicine* 200:129-135.

Kim SH, Frisina RD, Frisina DR. (2006). Effects of age on speech understanding in normal hearing listeners: Relationship between the auditory efferent system and speech intelligibility in noise. *Speech Communication* 48:855-862.

Koch DB, Downing M, Osberger MJ, Litvak L. (2007). Using current steering to increase spectral resolution in CII and HiRes 90K users. *Ear and Hearing* 28:38S-41S.

Koch DB, Osberger MJ, Segel P, Kessler D. (2004). HiResolutionTM and conventional sound processing in the HiResolutionTM bionic ear: using appropriate outcome measures to assess speech recognition ability. *Audiology and Neurotology* 9:214-223.

Koelewijn T, de Kluiver H, Shinn-Cunningham BG, Zekveld AA, Kramer SE. (2015). The pupil response reveals increased listening effort when it is difficult to focus attention. *Hearing Research* 323:81-90.

Koelewijn T, Versfeld NJ, Kramer SE. (2017). Effects of attention on the speech reception threshold and pupil response of people with impaired and normal hearing. *Hearing Research* 354:56-63.

Koelewijn T, Zekveld AA, Festen JM, Kramer SE. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear and Hearing* 33:291-300.

Kramer SE, Kapteyn TS, Festen JM, Kuik DJ. (1997). Assessing aspects of auditory handicap by means of pupil dilatation. *International Journal of Audiology* 36:155-164.

Kramer SE, Kapteyn TS, Houtgast T. (2006). Occupational performance: Comparing normally-hearing and hearing-impaired employees using the Amsterdam Checklist for Hearing and Work. *International Journal of Audiology* 45:503-512.

Laback B, Deutsch WA, Baumgartner WD. (2004). Coding of vowellike signals in cochlear implant listeners. *Journal of the Acoustical Society of America* 116:1208-1223.

Landsberger DM, Marozeau J, Mertens G, Van de Heyning P. (2018). The relationship between time and place coding with cochlear implants with long electrode arrays. *Journal of the Acoustical*

Society of America 144:EL509-EL514.

Larsby B, Hällgren M, Lyxell B, Arlinger S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology* 44:131-143.

Laszig R, Aschendorff A, Stecker M, Müller-Deile J, Maune S, Dillier N, (…) Doering W. (2004). Benefits of bilateral electrical stimulation with the nucleus cochlear implant in adults: 6-Month postoperative results. *Otology and Neurotology* 25:958-968.

Lenarz T. (2017). Cochlear implant–state of the art. *Laryngo- Rhino- Otologie* 96:S123-S151.

Liberman MC, Guinan JJ. (1998). Feedback control of the auditory periphery: Anti-masking effects of middle ear muscles vs. olivocochlear efferents. *Journal of Communication Disorders* 31:471-483.

Lilaonitkul W, Guinan JJ. (2009a). Human medial olivocochlear reflex: effects as functions of contralateral, ipsilateral, and bilateral elicitor bandwidths. *Journal of the Association for Research in Otolaryngology* 10:459-470.

Litovsky R, Parkinson A, Arcaroli J, Sammeth C. (2006). Simultaneous bilateral cochlear implantation in adults. *Ear and Hearing* 27:714-31.

Litovsky RY, Gordon K. (2016). Bilateral cochlear implants in children: Effects of auditory experience and deprivation on auditory perception. *Hearing Research* 338:76-87

Litovsky RY, Goupell MJ, Godar S, Grieco-Calub T, Jones GL, Garadat SN, (...) and Misurelli S. (2012). Studies on bilateral cochlear implants at the University of Wisconsin’s Binaural Hearing and Speech Laboratory. *Journal of the American Academy of Audiology* 23:476-494.

Litovsky RY, Parkinson A, Arcaroli J, Peters R, Lake J, Johnstone P, Yu G. (2004). Bilateral cochlear implants in adults and children. *Archives of Otolaryngology - Head and Neck Surgery* 130:648-655.

Loizou PC, 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. *Journal of the Acoustical Society of America* 125:372-383.

Loizou PC. (1999). Introduction to cochlear implant. *IEEE Engineering in Medical and Biology Magazine* 18:32-42.

Lopez-Poveda EA, Eustaquio-Martin A. (2014). Mimicking the unmasking effects of the medial olivo-cochlear efferent reflex with cochlear implants. *Journal of Acoustical Society of America* 135:2385.

Lopez-Poveda EA, Eustaquio-Martín A. (2018). Objective speech transmission improvements with a binaural cochlear implant sound-coding strategy inspired by the contralateral medial olivocochlear reflex. *Journal of the Acoustical Society of America* 143:2217-2231.

Lopez-Poveda EA, Eustaquio-Martin A, Stohl JS, Wolford RD, Schatzer R, Wilson BS. (2016b). A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Ear and Hearing* 37:138-148.

Lopez-Poveda EA, Eustaquio-Martín A, Stohl JS, Wolford RD, Schatzer R, Wilson BS. (2016a). Roles of the contralateral efferent reflex in hearing demonstrated with cochlear implants. *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing* 105-114.

Lopez-Poveda EA, Eustaquio-Martín A, Stohl JS, Wolford RD, Schatzer R, Gorospe JM, (…) and Wilson

BS. (2017). Intelligibility in speech maskers with a binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Hearing Research* 348:134-137.

Lopez-Poveda EA. (1996). The physical origin and physiological coding of pinna-based spectral cues. Doctoral Thesis. Loughborough University. United Kingdom.

Lopez-Poveda EA. (2014). Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech. *Frontiers in Neuroscience* 8:1-7.

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

Lopez-Poveda EA. (2018). Olivocochlear efferents in animals and humans: From anatomy to clinical relevance. *Frontiers in Neurology* 9:1-18.

Lorens A, Zgoda M, Obrycka A, Skarżynski H. (2010). Fine structure processing improves speech perception as well as objective and subjective benefits in pediatric MED-EL COMBI 40+ users. *International Journal of Pediatric Otorhinolaryngology* 74:1372-1378.

Lunner T. (2003). Cognitive function in relation to hearing aid use. *International Journal of Audiology* 42:49-58.

MacKeith NW, Coles RRA. (1971). Binaural advantages in hearing of speech. *Journal of Laryngology and Otology* 85:213-232.

Mackersie CL, Cones H. (2011). Subjective and psychophysiological indexes of listening effort in a competing-talker task. *Journal of the American Academy of Audiology* 22:113-122.

Mackersie CL, MacPhee IX, Heldt EW. (2015). Effects of hearing loss on heart-rate variability and skin conductance measured during sentence recognition in noise. *Ear and hearing* 36:145-154.

Majdak P, Goupell MJ, Laback B. (2011). Two-Dimensional Sound Localization in Cochlear Implantees. *Ear and Hearing* 32:198-208.

Majdak P, Masiero B, Fels J. (2013). Sound localization in individualized and non-individualized crosstalk cancellation systems. *Journal of the Acoustical Society of America* 133:2055-2068.

Marmel F, Marrufo-Pérez MI, Heeren J, Ewert S, Lopez-Poveda EA. (2018). Effect of sound level on virtual and free-field localization of brief sounds in the anterior median plane. *Hearing Research* 365:28-35.

McCoy SL, Tun PA, Cox LC, Colangelo M, Stewart RA, Wingfield A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology* 58:22-33.

McGarrigle R, Gustafson SJ, Hornsby BWY, Bess FH. (2019). Behavioral measures of listening effort in school-age children: Examining the effects of signal-to-noise ratio, hearing loss, and amplification. *Ear and Hearing* 40:381-392.

McGarrigle R, Munro KJ, Dawes P, Stewart AJ, Moore DR, Barry JG, 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." *International Journal of Audiology* 53:433-440.

McKay CM, McDermott HJ, Vandali A, and Clark GM. (1991). Preliminary results with a six spectral maxima speech processor for the University of Melbourne/Nucleus multiple electrode cochlear implant. *Journal of the Otolaryngological Society of Australia* 6:354-359.

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 in Hearing* 22:1-11.

Miles K, McMahon C, Boisvert I, Ibrahim R, de Lissa P, Graham P, Lyxell B. (2017). Objective assessment of listening effort: Coregistration of pupillometry and EEG. *Trends in Hearing* 21*:*1-13.

Mishra SK, Lutman ME. (2014). Top-down influences of the medial olivocochlear efferent system in speech perception in noise. *PLoS ONE 9*:e85756.

Misurelli SM, Litovsky RY. (2015). Spatial release from masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity. *Journal of the Acoustical Society of America* 138:319-331.

Møller AR. (2006). History of cochlear implants and auditory brainstem implants. *Advances in Oto-Rhino-Laryngology* 64:1-10.

Moore BCJ, Kolarik A, Stone MA, Lee Y-W. (2016). Evaluation of a method for enhancing interaural level differences at low frequencies. *Journal of the Acoustical Society of America* 140:2817-2828.

Moore BCJ. (2003). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology and Neurotology* 24:243-254.

Moore BCJ. (2007). *Cochlear hearing loss: physiological, psychological and technical issues*. John Wiley and Sons.

Müller J, Schon F, Helms J. (2002). Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system. *Ear and Hearing* 23:198-206.

Murugasu E, Russell IJ. (1996). The effect of efferent stimulation on basilar membrane displacement in the basal turn of the guinea pig cochlea. *Journal of Neuroscience* 16:325-332.

Nachtegaal J, Kuik DJ, Anema JR, Goverts ST, Festen JM, Kramer SE. (2009). Hearing status, need for recovery after work, and psychosocial work characteristics: Results from an internet-based national survey on hearing. *International Journal of Audiology* 48:684-691.

National Research Council. (2005). *Hearing loss: Determining Eligibility for Social Security Benefits*. Washington, DC: The National Academies Press.

NICE. (2009). Cochlear implants for children and adults with severe to profound deafness. *NICE Technology Appraisal Guidance* 166:1-41.

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

Nilsson M, Soli SD, Sullivan JA. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America* 95:1085-1099.

Nogueira W, Büchner A, Lenarz T, Edler B. (2005). A psychoacoustic 'NofM'-type speech coding strategy for cochlear implants. *EURASIP Journal of Applied Signal Processing* 18:3044-3059.

Nopp P, Schleich P, D'Haese P. (2004). Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants. *Ear and Hearing* 25*:*205-214.

Ohlenforst B, Zekveld AA, Jansma EP, 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 and Hearing* 38:267-281.

Ohlenforst B, Zekveld AA, Jansma EP, Wang Y, Naylor G, Lorens A, (…) and Kramer SE. (2017). Effects of hearing impairment and hearing aid amplification on listening effort: A systematic review. *Ear and Hearing* 38:267-281.

Pals C, Sarampalis A, Başkent D. (2012). Listening effort with cochlear implant simulations. *Journal of Speech, Language, and Hearing Research* 56:1075-1084.

Pals C, Sarampalis A, van Rijn H, Başkent D. (2015). Validation of a simple response-time measure of listening effort. *Journal of the Acoustical Society of America* 138:187-192.

Peelle JE. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing* 39*:*204-214.

Perreau AE, Wu YH, Tatge B, Irwin D, Corts D. (2017). Listening effort measured in adults with normal hearing and cochlear implants. *Journal of the American Academy of Audiology* 28: 685-697.

Peterson NR, Pisoni DB, Miyamoto RT. (2010). Cochlear implants and spoken language processing abilities: Review and assessment of the literature. *Restorative Neurology and Neuroscience* 28:237-250.

Pichora-Fuller MK, Kramer SE, Eckert MA, Edwards B, Hornsby BWY, Humes LE, (…) and Wingfield A. (2016). Hearing impairment and cognitive energy. *Ear and Hearing* 37:5S-27S.

Pichora-Fuller MK, Singh G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification* 10:29-59.

Picou EM, Gordon J, Ricketts TA. (2016). The effects of noise and reverberation on listening effort in adults with normal hearing. *Ear and Hearing* 37:1-13.

Picou EM, Moore TM, Ricketts TA. (2017). The effects of directional processing on objective and subjective listening effort. *Journal of Speech, Language, and Hearing Research* 60:199-211.

Picou EM, Ricketts TA, Hornsby BW. (2013). How hearing aids, background noise, and visual cues influence objective listening effort. *Ear and Hearing* 34:e52-e64.

Picou EM, Ricketts TA. (2014). The effect of changing the secondary task in dual-task paradigms for measuring listening effort. *Ear and Hearing* 35:611-622.

Potts WB, Ramanna L, Perry T, Long CJ. (2019). Improving localization and speech reception in noise for bilateral cochlear implant recipients. *Trends in Hearing* 23:1-18.

Robles L, Ruggero MA. (2001). Mechanics of the mammalian cochlea. *Physiological Reviews* 81:1305-1352.

Ruggero MS, Robles L, Rich NC (1992). Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression. *Journal of Neurophysiology* 68:1087-1099.

Sahin MI, Sagers JE, Stankovic KM. (2017). Cochlear implantation: Vast unmet need to address deafned globaly. *Otolology and Neurotology* 38:786-787.

Sarampalis A, Kalluri S, Edwards B, Hafter E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech, Language, and Hearing Research* 52:1230-1240.

Sarant J, Harris D, Bennet L, Bant S. (2014). Bilateral versus unilateral cochlear implants in children.

Ear and Hearing 35:396-409.

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 and Hearing* 25:197-204.

Seeber BU, Fastl H. (2008). Localization cues with bilateral cochlear implants. *Journal of the Acoustical Society of America* 123:1030-1042.

Seeber BU, Baumann U, Fastl H. (2004). Localization ability with bimodal hearing aids and bilateral cochlear implants. *Journal of the Acoustical Society of America* 116:1698-1709.

Shannon RV, Fu Q-J, Galvin J, Friesen L. (2003). Speech perception with cochlear implants. Chapter 8 in Auditory Prostheses, F-G Zeng, AN Popper, and RR Fay, Eds. Springer Handbook of Auditory Research, Volume X, 213:1-43.

Simmons DD. (2002). Development of the inner ear efferent system across vertebrate species. *Journal of Neurobiology* 53:228-250.

Smeds K, Wolters F, Rung M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology* 26:183-196.

Smith DW, Turner DA, Henson MM. (2000). Psychophysical correlates of contralateral efferent suppression. I. The role of the medial olivocochlear system in "central masking" in nonhuman primates. *Journal of the Acoustical Society of America* 107:933-941.

Smulders YE, Van Zon A, Stegeman I, Rinia AB, Van Zanten GA, Stokroos RJ, (…) and Grolman W. (2016). Comparison of bilateral and unilateral cochlear implantation in adults a randomized clinical trial. *JAMA Otolaryngology-Head and Neck Surgery* 142:249-256.

Spencer NJ, Tillery KH, Brown CA. (2019). The effects of dynamic-range automatic gain control on sentence intelligibility with a speech masker in simulated cochlear implant listening. *Ear and Hearing* 40:710-724.

Stevens G, Flaxman S, Brunskill E, Mascarenhas M, Mathers CD, Finucane M. (2013). Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. *European Journal of Public Health* 23*:*146-152.

Stöbich B, Zierhofer CM, Hochmair ES. (1999). Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy. *Ear and Hearing* 20:104-116.

Studebaker GA, McDaniel DM, Sherbecoe RL. (1995). Evaluating relative speech recognition performance using the proficiency factor and rationalized arcsine differences. *Journal of the American Academy of Audiology* 6:173-182.

Studebaker GA. (1985). A "rationalized" arcsine transform. *Journal of Speech, Language, and Hearing Research* 28:455-462.

Svirsky MA, Teoh SW, Neuburger H. (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiology and Neuro-Otology* 9:224-233.

Swaminathan J, Mason CR, Streeter TM, Best V, Roverud E, Kidd G, Jr (2016). Role of binaural temporal fine structure and envelope cues in cocktail-party listening. *Journal of Neuroscience* 36:8250-8257.

Taal CH, Hendriks RC, Heusdens R, Jensen J. (2010). On predicting the difference in intelligibility before and after single-channel noise reduction. *International Workshop, Acoustic Echo Noise*

Control (Tel Aviv, Israel).

Taal CH, Hendriks RC, Heusdens R, Jensen J. (2011). An algorithm for intelligibility prediction of time-frequency weighted noisy speech. *IEEE Transactions on Audio, Speech and Language Processing* 19:2125-2136.

Theelen-van den Hoek FL, Boymans M, van Dijk B, Dreschler WA. (2016). Adjustments of the amplitude mapping function: Sensitivity of cochlear implant users and effects on subjective preference and speech recognition. *International Journal of Audiology* 55:674-687.

Tong YC, Blamey PJ, Dowell RC, Clark GM. (1983). Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant. *Journal of the Acoustical Society of America* 74:73-80.

Tun PA, McCoy S, Wingfield A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychology and Aging* 24:761-766.

Tyler RS, Gantz BJ, Rubinstein JT, Wilson BS, Parkinson AJ, Wolaver A, (…) and Lowder MW. (2002). Three-month results with bilateral cochlear implants. *Ear and Hearing* 23:80S-89S.

Tyler RS, Noble W, Dunn C, Witt S. (2006). Some benefits and limitations of binaural cochlear implants and our ability to measure them. *International Journal of Audiology* 45:113-119.

van den Tillaart-Haverkate M, de Ronde-Brons I, Dreschler WA, Houben R. (2017). The influence of noise reduction on speech ntelligibility, response times to speech, and perceived listening effort in normal-hearing listeners. *Trends in Hearing* 21:1-13.

van Hoesel RJM, Tyler RS. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *Journal of the Acoustical Society of America* 113:1617-1630.

van Hoesel RJM. (2004). Exploring the benefits of bilateral cochlear implants. *Audiology and Neuro-Otology* 9:234-246.

Verschuur CA, Lutman ME, Ramsden R, Greenham P, O'Driscoll M. (2005). Auditory localization abilities in bilateral cochlear implant recipients. *Otology and Neurotology* 26:965-971.

von Békésy G, Wever EG. (1960). *Experiments in hearing* (Vol. 8). New York: McGraw-Hill.

Warr WB, Guinan JJ (1979). Efferent innervation in the organ of Corti: two separate systems. *Brain Research* 173:152-155.

Wenzel EM, Arruda M, Kistler DJ, Wightman FL. (1993). Localization using nonindividualized head-related transfer functions. *Journal of the Acoustical Society of America* 94:111-123.

Wild CJ, Yusuf A, Wilson DE, Peelle JE, Davis MH, Johnsrude IS. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *Journal of Neuroscience* 32:14010-14021.

Wilson BS, Dorman MF. (2007). The surprising performance of present-day cochlear implants. *IEEE Transactions on Biomedical Engineering* 54:969-972.

Wilson BS, Dorman MF. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research* 242:1-2.

Wilson BS, Dorman MF. (2018). *A brief history of the cochlear implant and related treatments. Neuromodulation* (Second Edi). Elsevier Ltd. Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM. (1991). Better speech recognition with cochlear implants. *Nature* 352:236-238.

Wilson BS, Finley CC, Weber BA, White MW, Farmer Jr JC, Wolford RD, (...) and Schindler RA. (1988). Comparative studies of speech processing strategies for cochlear implants. *Laryngoscope* 98:1069-1077.

Wilson BS, Schatzer R, Lopez-Poveda EA, Sun X, Lawson DT, Wolford RD. (2005). Two new directions in speech processor design for cochlear implants. *Ear and Hearing* 26:73s-81s.

Wilson, B. S. (2013). Toward better representations of sound with cochlear implants. *Nature Medicine* 19:1245-1248.

Wilson BS. (2015). Getting a decent (but sparse) signal to the brain for users of cochlear implants. *Hearing Research* 322:24-38.

Wilson BS. (2017). The cochlear implant and possibilities for narrowing the remaining gaps between prosthetic and normal hearing. *World Journal of Otorhinolaryngology - Head and Neck Surgery* 3:200-210.

Wilson, B. S. (2017). The modern cochlear implant: a triumph of biomedical engineering and the first substantial restoration of human sense using a medical intervention. *IEEE pulse* 8: 29-32.

Winn M. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing* 20:1-17.

Winslow RL, Sachs MB. (1988). Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. *Hearing Research* 35*:*165-189.

World Health Organisation. (2018). *Addressing the rising prevalence of hearing loss*.

Wouters J, McDermott HJ, Francart T. (2015). Sound coding in cochlear implants: From electric pulses to hearing. *IEEE Signal Processing Magazine* 32:67-80.

Xu K, Willis S, Gopen Q, Fu Q-J. (2020) Effects of Spectral Resolution and Frequency Mismatch on Speech Understanding and Spatial Release From Masking in Simulated Bilateral Cochlear Implants. *Ear and Hearing* 41:1362-1371.

Yoshinaga-Itano C, Sedey AL, Coulter DK, Mehl AL. (1998). Language of early- and later-identified children with hearing loss. *Pediatrics* 102:1161-1171.

Yost WA, Dye Jr RH. (1988). Discrimination of interaural differences of level as a function of frequency. *Journal of the Acoustical Society of America* 83:1846-1851.

Zekveld AA, Kramer SE, Festen JM. (2010). Pupil response as an indication of effortful listening: the influence of sentence intelligibility. *Ear and Hearing* 31:480-490.

Zekveld AA, Kramer SE, Festen JM. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and Hearing 32*:498-510.

Zeng FG, Rebscher S, Harrison W, Sun X, Feng H. (2008). Cochlear implants: System design, integration, and evaluation. *IEEE reviews in biomedical engineering* 1:115-142.

Zeng FG, Rebscher SJ, Fu QJ, Chen H, Sun X, Yin L, (...) and Yang B. (2015). Development and evaluation of the Nurotron 26-electrode cochlear implant system. *Hearing research* 322:188-199.

Zeng FG. (2004). Trends in cochlear implants. *Trends in Amplification* 8:1-34.

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

WO 01/13991 A1.

Appendix 1.

REPRINTS OF PUBLISHED ARTICLES

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;XX;00–00)

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

---

[1]Laboratorio de Audición Computacional y Psicoacústica, Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Salamanca, Spain; [2]Grupo de Audiología, Instituto de Investigación Biomédica de Salamanca, Universidad de Salamanca, Salamanca, Spain; [3]Departamento de Cirugía, Facultad de Medicina, Universidad de Salamanca, Salamanca, Spain; [4]Unidad de Foniatría, Logopedia y Audiología, Servicio de Otorrinolaringología, Hospital Universitario de Salamanca, Salamanca, Spain; [5]Servicio de Otorrinolaringología, Hospital Universitario Ramón y Cajal, Madrid, Spain; [6]MED-EL GmbH, Innsbruck, Austria; and [7]North American Research Laboratory, MED-EL Corporation, Durham, North Carolina, USA.

0196/0202/2020/XXXX-00/0 • Ear & Hearing • Copyright © 2020 The Authors. Ear & Hearing is published on behalf of
the American Auditory Society, by Wolters Kluwer Health, Inc. • Printed in the U.S.A.

<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):

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

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:

$$E' = E \cdot \sqrt{\frac{BW_{\mathrm{ref}}}{BW}}, \tag{2}$$

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.

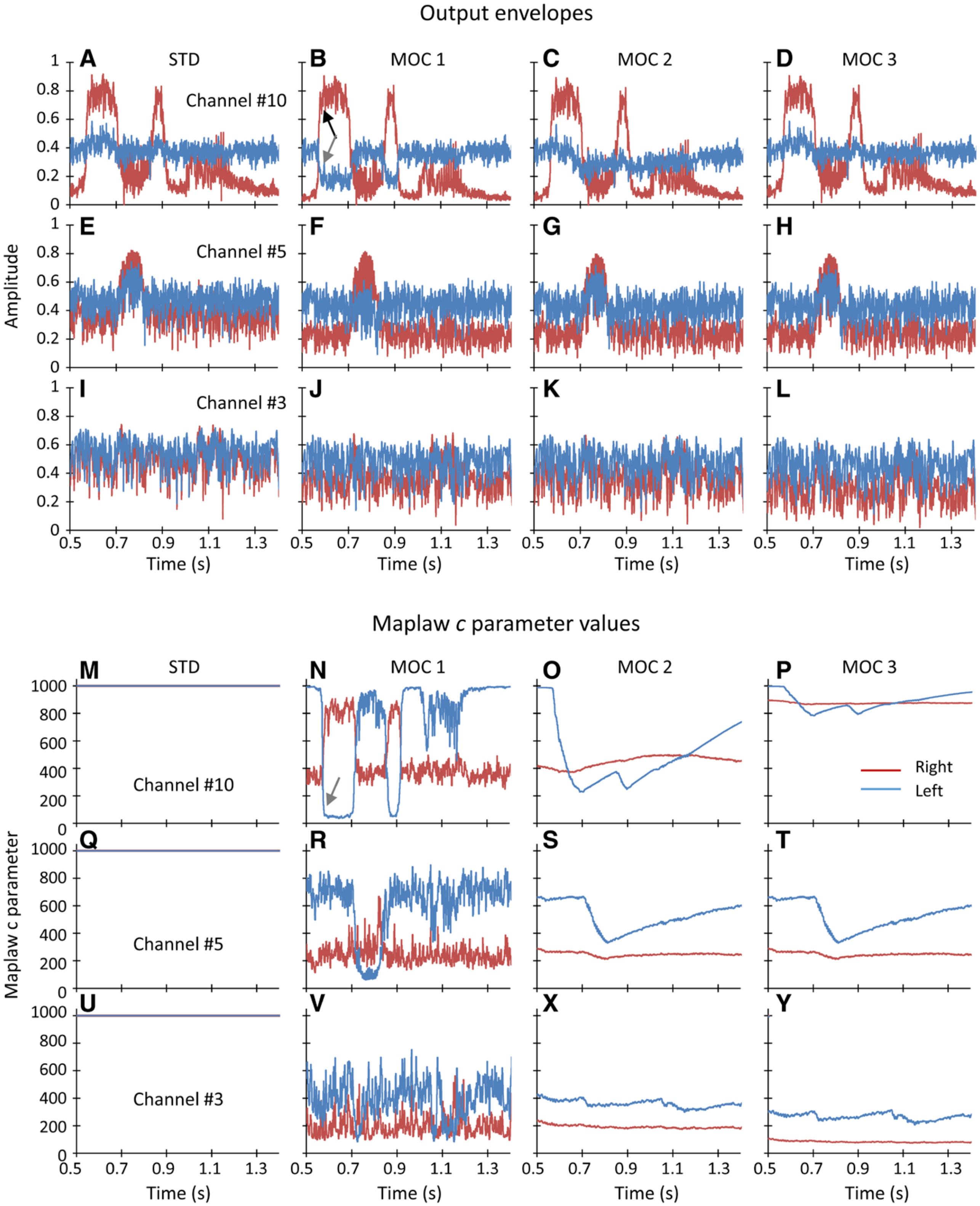

Fig. 1. Example compressed envelopes (A–L) and maplaw values (M–Y) for STD, MOC1, MOC2, and MOC3 strategies with 10 frequency channels. Data are shown only for three channels: channel number 3 (bottom row), channel number 5 (E–H and Q–T), and channel number 10 (top row) with center frequencies of 501, 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

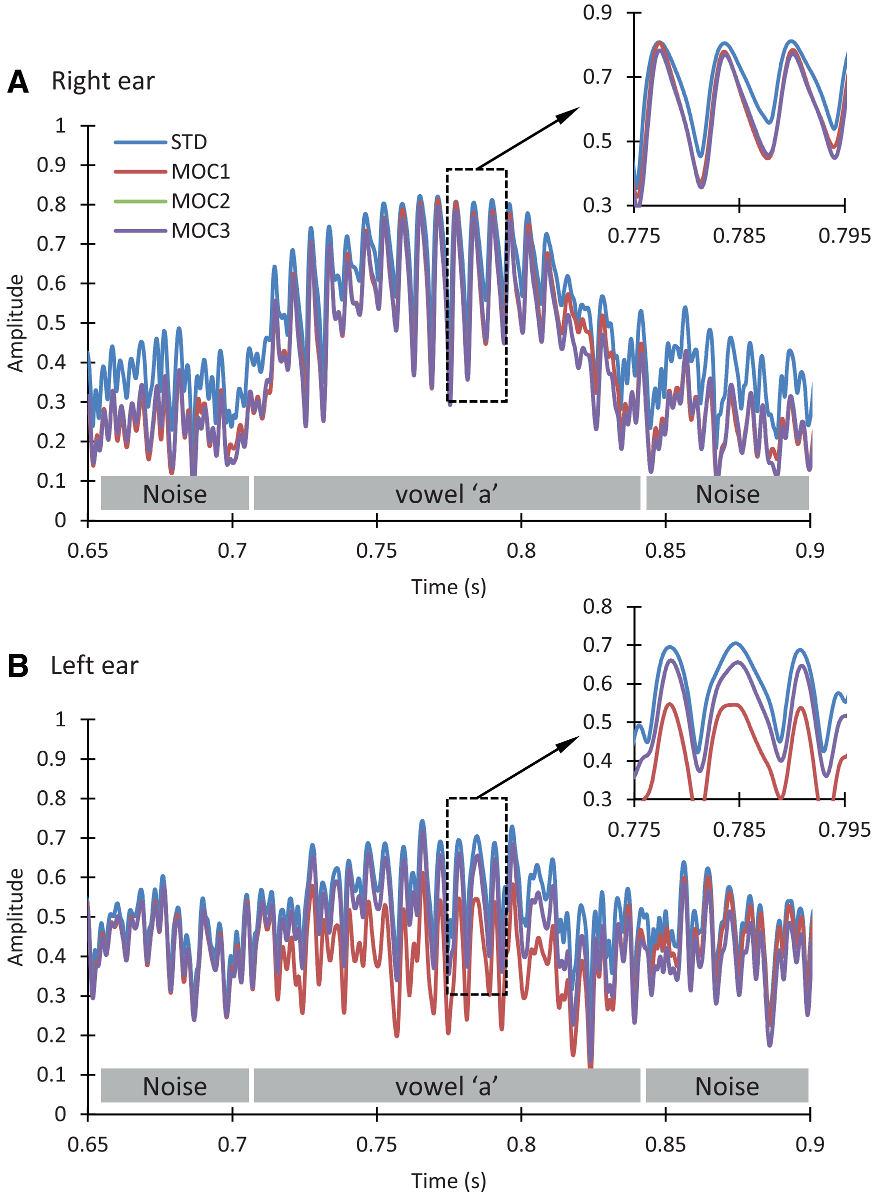

Fig. 2. Zoomed-in view of the compressed envelopes for channel number 5 shown in Fig. 1. Each panel shows envelopes for the STD, MOC1, MOC2, and MOC3 strategies. Envelopes were identical for the MOC2 and MOC3, hence the overlap between corresponding traces. The gray rectangles near the abscissae depict 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

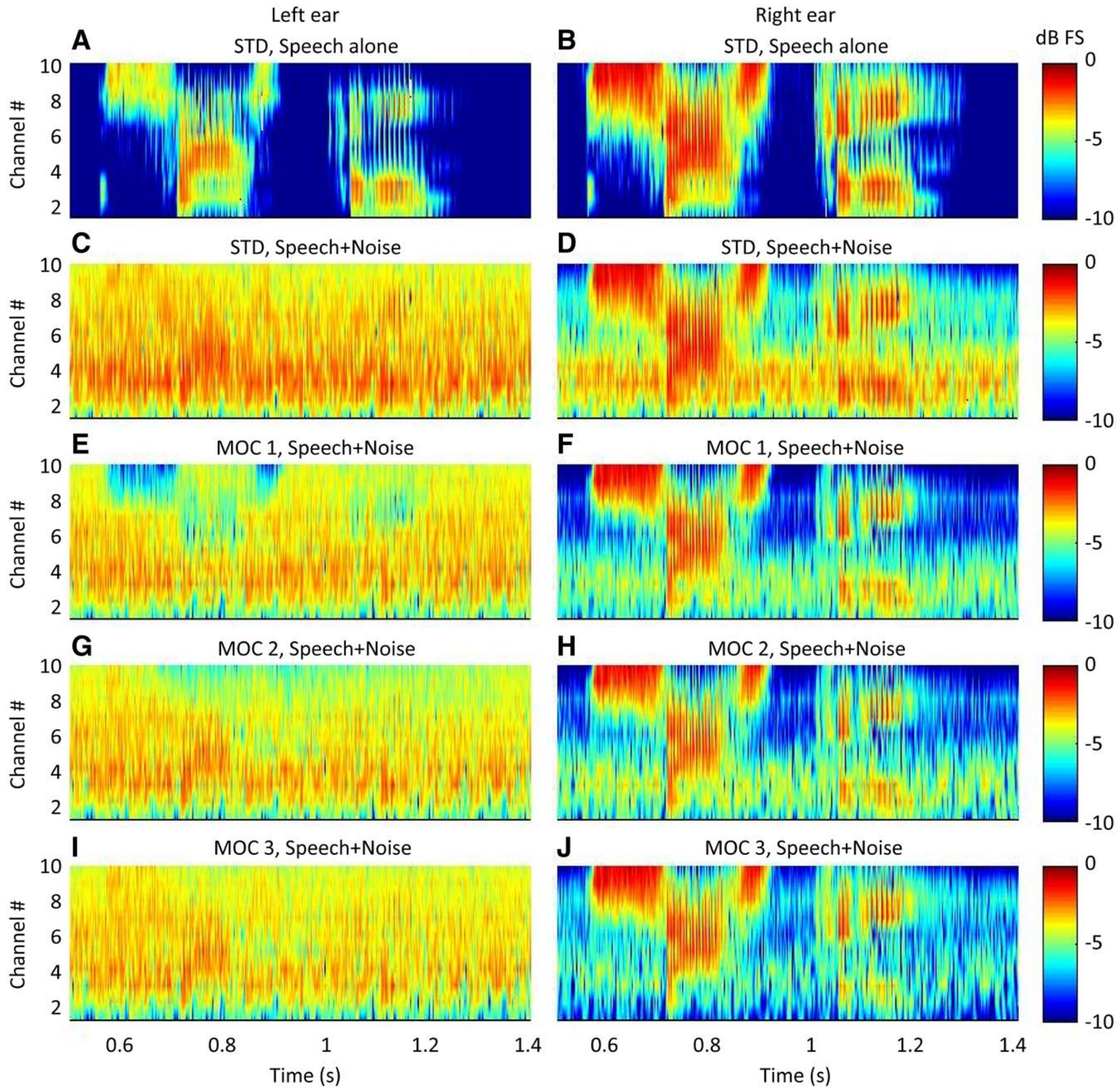

Fig. 3. Output envelopes for STD, MOC1, MOC2, and MOC3 processors with 10 frequency channels. The stimulus was as in Fig. 1. Each panel shows envelopes at the output of the maplaw as a function of frequency channel number and time. Color illustrates amplitude in units of dB FS, and spatial smoothing 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:

$$\mathrm{SRT_{benefit}}\,[\mathrm{dB}] = \mathrm{SRT_{STD}}\,[\mathrm{dB\ SNR}] - \mathrm{SRT_{MOC}}\,[\mathrm{dB\ SNR}] \qquad (3)$$

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

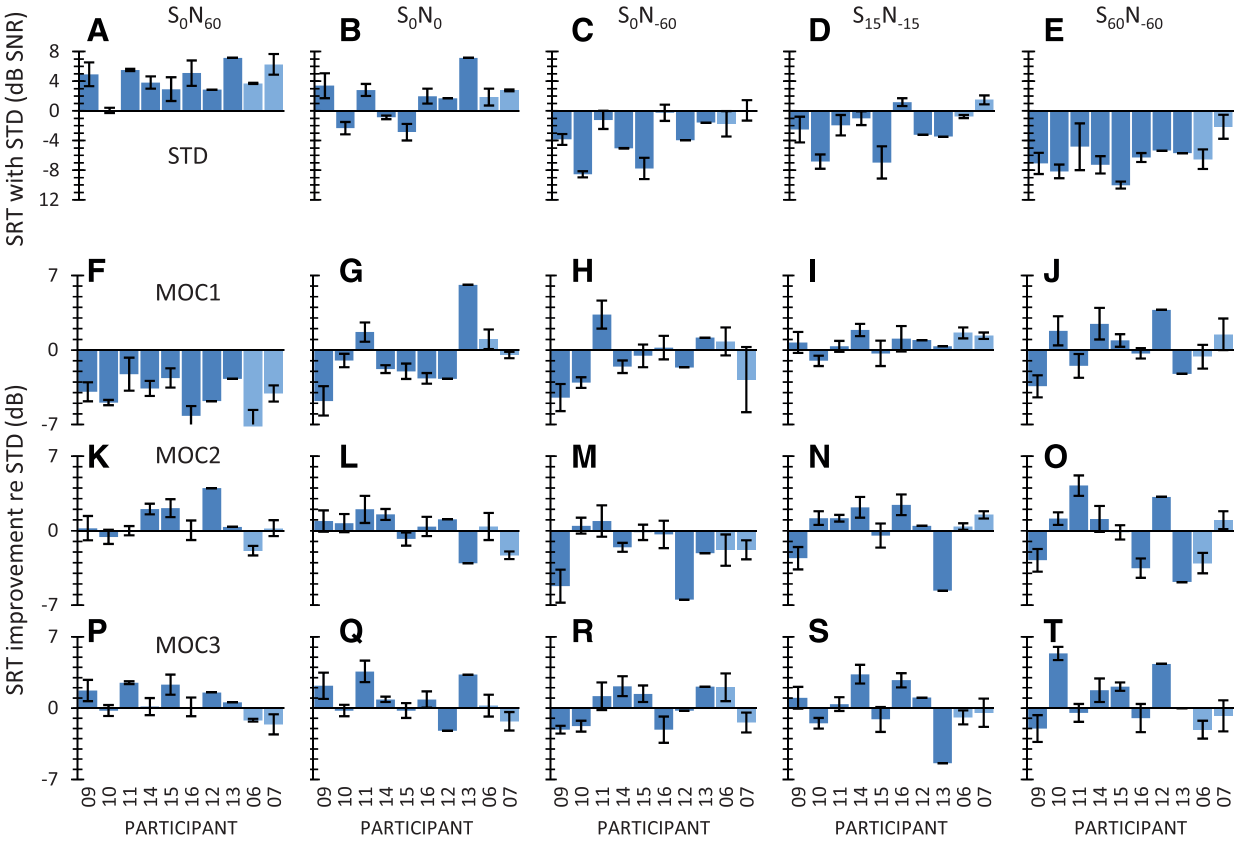

Fig. 4. Intelligibility in unilateral listening for individual participants. Row 1 (panels A to E), SRTs for the STD strategy. Each panel is for a different spatial configuration of the speech and noise sources, as indicated at the top. Rows 2 to 4 (panels F to T), SRT improvement relative to the STD 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.

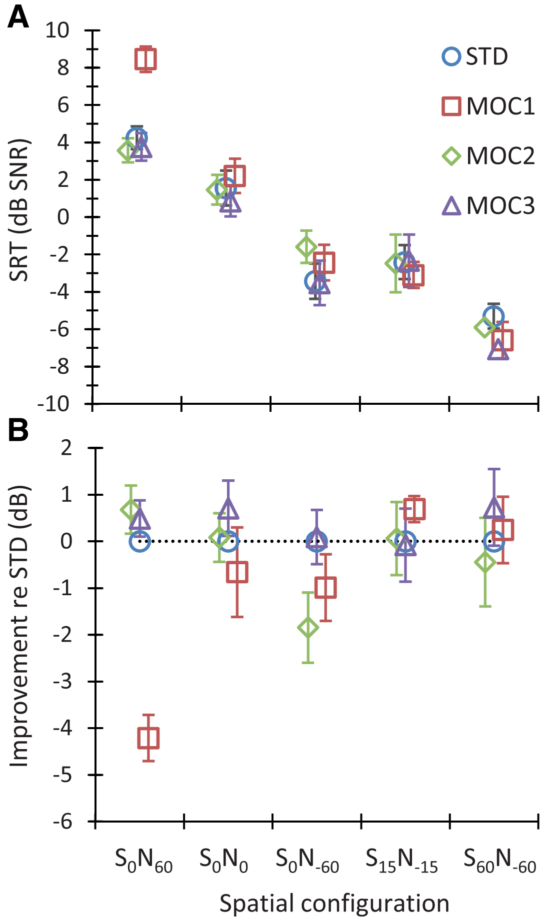

Fig. 5. Group mean intelligibility scores in unilateral listening. A, Mean SRTs for each strategy (as indicated by the inset) and spatial configuration (as indicated in the abscissa). Each point is the mean for eight bilateral and two unilateral CI users. B, Mean SRT improvement for the MOC strategies 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

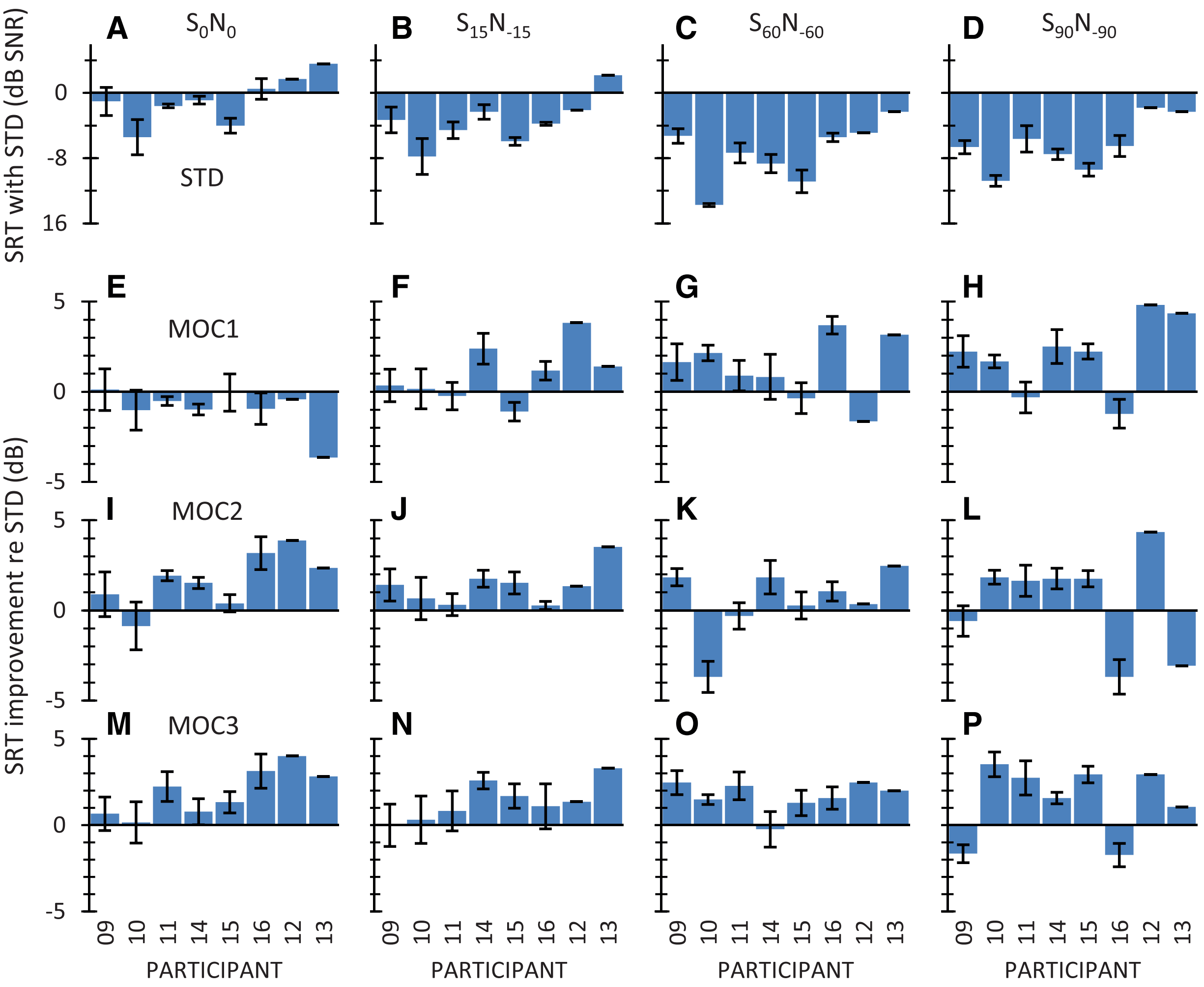

Fig. 6. Intelligibility in bilateral listening for individual participants. The layout is the same as Fig. 4. MOC indicates medial olivocochlear; MOC1, original fast MOC strategy; MOC2, slower MOC strategy; MOC3, slower MOC strategy with comparatively greater contralateral inhibition in the lower-frequency than in the higher-frequency channels; N, noise; S, speech; SRT, speech reception threshold; STD, standard.

[p(MOC1 versus MOC2) = 1.00; p(MOC1 versus MOC3) = 0.29; p(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(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

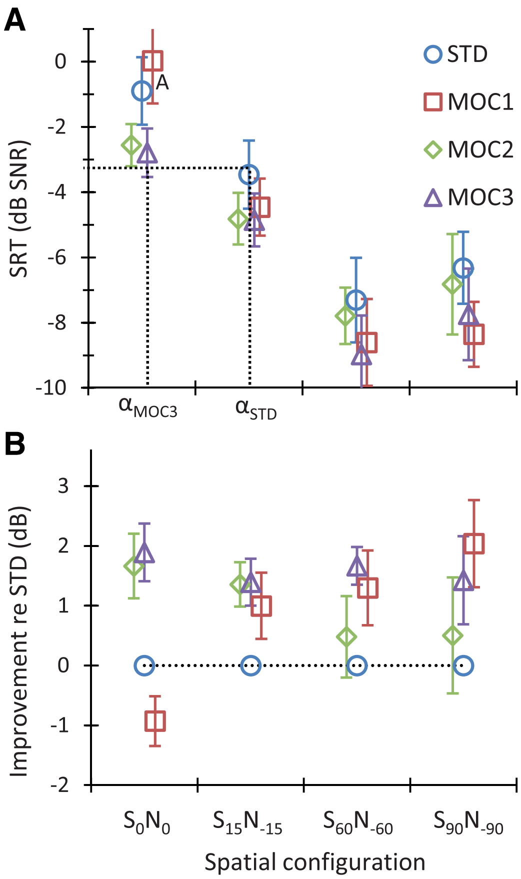

Fig. 7. Group mean intelligibility scores in bilateral listening. Each point is the mean for eight bilateral CI users. The layout is the same as Fig. 5. The dotted lines in panel A illustrate that at a fixed SNR of about −3 dB, the angular separation between the speech and noise source (α) to achieve 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

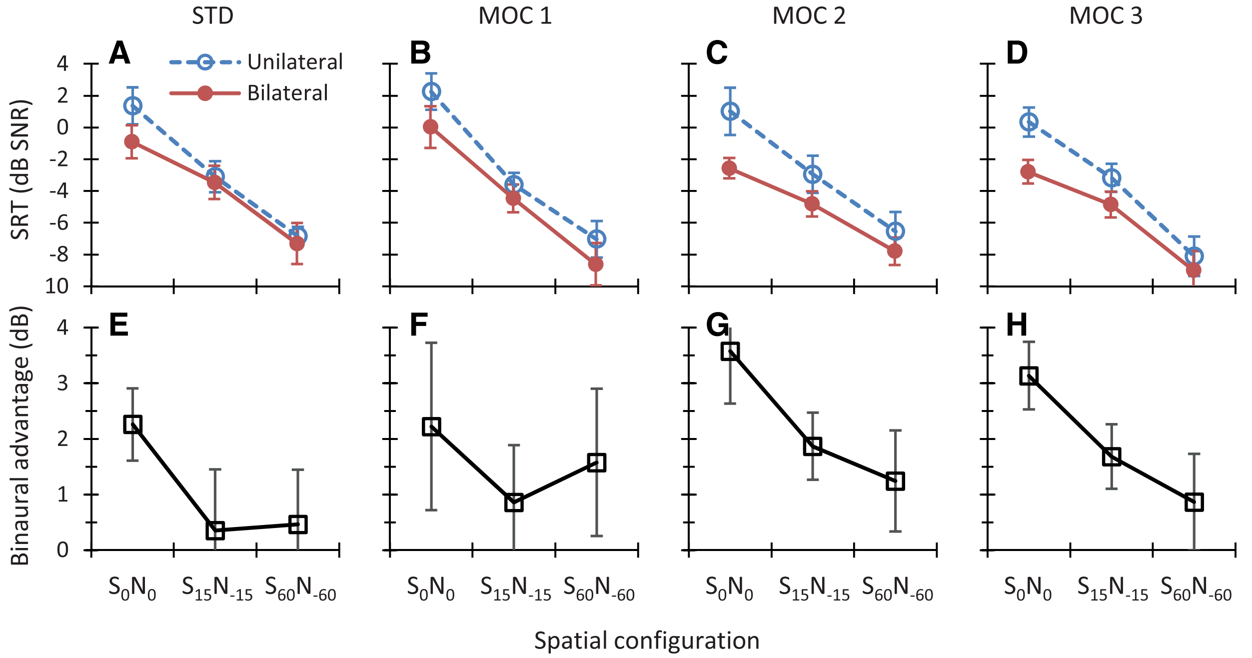

Fig. 8. Top, Group mean SRTs in unilateral and bilateral listening. Each panel is for a different strategy, as indicated at the top of the panel. Bottom, Mean binaural advantage calculated as the difference in mean SRT for unilateral listening minus bilateral listening. Positive values indicate better (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	$F(1,7) = 2.78, p = 0.139$	**$F(2,14) = 143.96, p < 0.001$**	$F(2,14) = 1.57, p = 0.240$
MOC1	8	$F(1,7) = 2.89, p = 0.130$	**$F(2,14) = 106.22, p < 0.001$**	$F(2,14) = 0.36, p = 0.700$
MOC2	8	**$F(1,7) = 10.36, p = 0.014$**	**$F(2,14) = 97.28, p < 0.001$**	**$F(2,14) = 4.32, p = 0.034$**
MOC3	8	**$F(1,7) = 20.22, p = 0.003$**	**$F(2,14) = 88.06, p < 0.001$**	$F(2,14) = 2.86, p = 0.091$

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

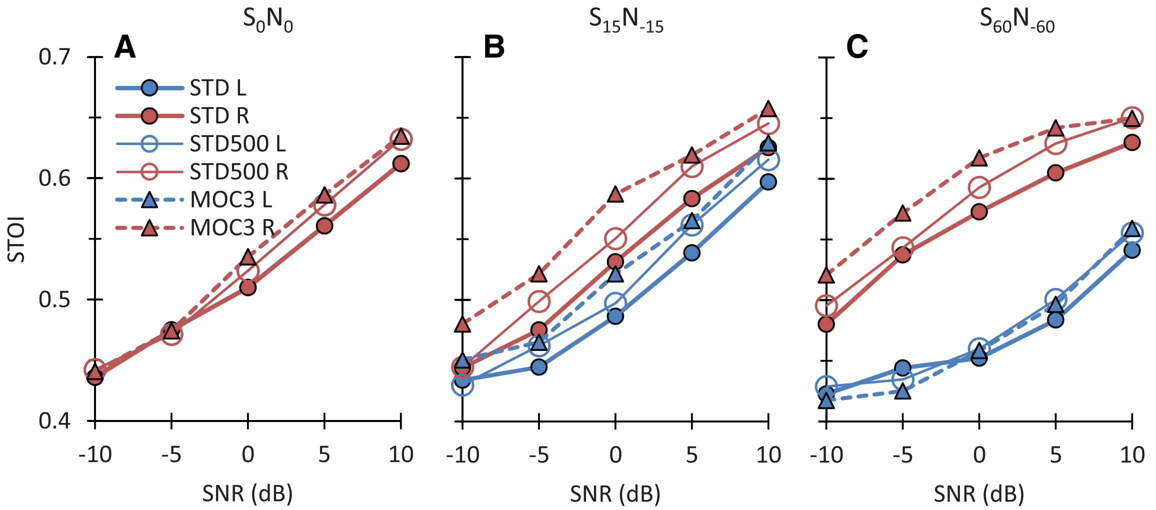

Fig. 9. Comparison of STOI scores for the present STD and MOC3 strategies against scores for a STD strategy with *c* = 500, the value typically used by the participants in their clinical audio processors. Each panel shows scores for the left (L) and right (R) ears (blue and red traces, respectively) 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.

ACKNOWLEDGMENTS

We thank Dr. Otto Peter from the University of Innsbruck for equipment and technical support. E.A.L.-P. conceived the MOC processors and the study. E.A.L.-P., A.E.-M., J.S.S., R.S., and P.N. designed the tested conditions. A.E.-M. and J.S.S. implemented the sound processors and testing tools. M.J.F., A.E.-M., and J.M.G. collected the data. A.E.-M. analyzed the data. J.M.G., R.P.L., and M.A.G.R. recruited the participants. E.A.L.-P., A.E.-M., and M.J.F. wrote the manuscript. All authors revised and edited the manuscript.

Work supported by MED-EL GmbH, the Spanish Ministry of Economy and Competitiveness (grant BFU2015-65376-P), and European Regional Development Funds.

Portions of this article were presented at the 41 Midwinter Meeting of the Association for Research in Otolaryngology, February 9–14, 2018, San Diego, California.

The authors have no conflicts of interest to disclose.

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

Received February 14, 2019; accepted March 24, 2020.

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.

Hearing Research 379 (2019) 103–116

Contents lists available at ScienceDirect

Hearing Research

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

Research Paper

Lateralization of virtual sound sources with a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex

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

[a] *Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Salamanca, Spain*
[b] *Instituto de Investigación Biomédica de Salamanca, Universidad de Salamanca, Salamanca, Spain*
[c] *Departamento de Cirugía, Facultad de Medicina, Universidad de Salamanca, Salamanca, Spain*
[d] *USA Laboratory, MED-EL GmbH, Innsbruck, Austria*
[e] *Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA*
[f] *MED-EL GmbH, Innsbruck, Austria*
[g] *Servicio de Otorrinolaringología, Hospital Universitario de Salamanca, Salamanca, Spain*
[h] *Servicio de Otorrinolaringología, Hospital Ramón y Cajal, Madrid, Spain*
[i] *School of Medicine and Pratt School of Engineering, Duke University, Durham, NC, USA*

ARTICLE INFO

Article history:
Received 16 November 2018
Received in revised form
30 April 2019
Accepted 17 May 2019
Available online 21 May 2019

Keywords:
Bilateral cochlear implant
Binaural processing
Binaural hearing
Olivocochlear efferents
MOC strategy
ILD
Sound localization

ABSTRACT

Many users of bilateral cochlear implants (BiCIs) localize sound sources less accurately than do people with normal hearing. This may be partly due to using two independently functioning CIs with fixed compression, which distorts and/or reduces interaural level differences (ILDs). Here, we investigate the potential benefits of using binaurally coupled, dynamic compression inspired by the medial olivocochlear reflex; an approach termed “the MOC strategy” (Lopez-Poveda et al., 2016, Ear Hear 37:e138-e148). Twelve BiCI users were asked to localize wideband (125–6000 Hz) noise tokens in a virtual horizontal plane. Stimuli were processed through a standard (STD) sound processing strategy (i.e., involving two independently functioning sound processors with fixed compression) and three different implementations of the MOC strategy: one with fast (MOC1) and two with slower contralateral control of compression (MOC2 and MOC3). The MOC1 and MOC2 strategies had effectively greater inhibition in the higher than in the lower frequency channels, while the MOC3 strategy had slightly greater inhibition in the lower than in the higher frequency channels. Localization was most accurate with the MOC1 strategy, presumably because it provided the largest and less ambiguous ILDs. The angle error improved slightly from 25.3° with the STD strategy to 22.7° with the MOC1 strategy. The improvement in localization ability over the STD strategy disappeared when the contralateral control of compression was made slower, presumably because stimuli were too short (200 ms) for the slower contralateral inhibition to enhance ILDs. Results suggest that some MOC implementations hold promise for improving not only speech-in-noise intelligibility, as shown elsewhere, but also sound source lateralization.

1. Introduction

Compared to unilateral cochlear implants (CIs), bilateral CIs (BiCIs) improve sound-source localization accuracy in the horizontal plane (e.g., Nopp et al., 2004; Seeber et al., 2004; van Hoesel, 2004; for a review, read the Introduction of Jones et al., 2014). Bilateral stimulation alone, however, is not enough to restore normal sound localization abilities. Indeed, most BiCI users show poorer localization scores for sound sources in the horizonal plane than do people with normal hearing (e.g., Seeber et al., 2004; Grantham et al., 2007; Majdak et al., 2011; Dorman et al., 2016).

---

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

https://doi.org/10.1016/j.heares.2019.05.004
0378-5955/© 2019 Elsevier B.V. All rights reserved.

This difference in performance is probably due to numerous factors, including the duration of auditory deprivation, the physiological status of the auditory nerve at the time of implantation, potential asymmetries in the age of cochlear implantation, and the (in) sensitivity of BiCI users to relevant acoustic information for sound localization (Kan and Litovsky, 2015).

In addition to the factors just listed, inadequate coding of binaural cues by the CI audio processors likely contributes to the poorer localization performance of BiCI users compared to normal-hearing listeners. The use of independently functioning devices, different number of frequency channels, and/or different rates of pulsatile electrical stimulation across the ears can distort and/or degrade interaural level difference (ILD) and interaural time difference (ITD) localization cues (Kan and Litovsky, 2015). However, even when the audio processors allow tight control over fine-structure ITDs, the best performers among BiCI users rely mostly on ILDs and much less so (or not at all) on ITDs to judge the location of sound sources in the horizontal plane (Dorman et al., 2016; Laback et al., 2004; Litovsky and Gordon, 2016; Seeber and Fastl, 2008). Hence, it seems important to faithfully encode ILDs to improve the localization performance of BiCI users.

One specific factor that can potentially degrade the coding of ILD cues is the use of independent compression in the two audio processors of a BiCI user. The CI audio processor typically includes a two-stage compression design to accommodate a broad range of acoustic pressure into a much narrower range of electrical current (Zeng, 2004). The first stage is a broadband automatic gain control (AGC) (Boyle et al., 2009; Stöbich et al., 1999). The AGC is placed at the front-end of processing and serves to narrow the broad range of ‘loudness’ fluctuations that occur naturally in the acoustic environment. The second compression stage is the acoustic-to-electric map. This map is placed at the back-end in each frequency channel of processing and serves to map the range of acoustic pressure into a narrower range of electrical current (e.g., Fu and Shannon, 1998; Wilson et al., 1991). The current standard (STD) is for BiCI users to wear two audio processors that function independently from each other. As illustrated in schematic form in Fig. 1A, the application of independent AGC and/or acoustic-to-electric maps to the two ears can compress (reduce) the head-shadow ILDs and thus hinder the localization of sound sources in the horizontal plane (e.g., Dorman et al., 2014; Ricketts et al., 2006; Wiggins and Seeber, 2011). Indeed, BiCI users can localize sounds more accurately with binaurally linked rather than with independent AGC in their two devices (Potts et al., 2019).

A binaural CI sound processing strategy has been recently proposed that uses dynamic (time-varying), binaurally *coupled* back-end compression inspired by the inhibitory effect of the contralateral medial olivocochlear (MOC) reflex on basilar membrane responses (Lopez-Poveda, 2015). The coupling is such that the greater the amplitude at the output of every frequency channel of processing in an audio processor, the more linear the back-end compression (or acoustic-to-electric map) in the corresponding frequency channel of the contralateral audio processor. The MOC strategy was intended to mimic the potential antimasking effects of the contralateral MOC reflex for speech-in-noise intelligibility with CIs. Indeed, the MOC strategy can improve the intelligibility of speech in competition with steady-state noise (Lopez-Poveda et al., 2016a) and single-talker interferers (Lopez-Poveda et al., 2017). Incidentally, however, by using binaurally coupled back-end compression, the MOC strategy can also enhance the head-shadow ILDs in each frequency channel of processing (see Fig. 2 in Lopez-Poveda et al., 2016a), thus the overall ILDs, relative to that available with two independently functioning processors (the STD approach). Fig. 1B illustrates the mechanism in schematic form. Insofar as BiCI users rely mostly on ILD cues for localization, it seems possible that sound source localization in the horizontal plane may be better with the MOC than with the STD strategy. The main aim of the present study was to investigate this possibility using virtual acoustic stimuli.

On the other hand, the ILDs delivered by the MOC strategy depend on the amount of contralateral inhibition of compression, which can be set using parameters. A second aim was to compare sound lateralization performance with various implementations of the MOC strategy designed to reflect more or less realistically the inhibitory characteristics of the natural contralateral MOC reflex (Lopez-Poveda and Eustaquio-Martín, 2018).

2. Methods

2.1. Participants

Twelve users of bilateral MED-EL CIs participated in the study (Table 1). Three of them were tested at the MED-EL US Laboratory (North Carolina, USA) and nine of them were tested at the

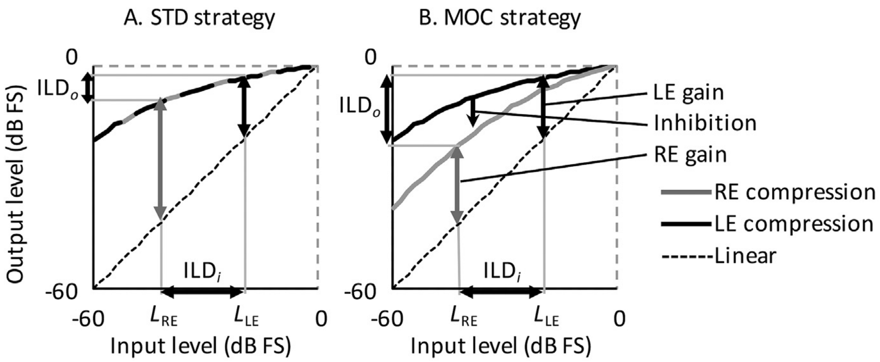

Fig. 1. Schematic interaural level difference at the output ($\mathrm{ILD_o}$) for a given interaural level difference at the input ($\mathrm{ILD_i}$) with two different back-end compression schemes. **A.** With equal and independently functioning compressors at the two ears (STD strategy). **B.** With binaurally coupled compressors as in the MOC strategy. Each panel illustrates the $\mathrm{ILD_o}$ (double headed arrow on the ordinate) for a hypothetical sound source located in the free field on the left side of the head; i.e., when the stimulus (input) level is greater on the left than on the right ear ($L_{\mathrm{LE}} > L_{\mathrm{RE}}$). With the STD strategy (**A**), the compression function would be identical for the right and the left ear, hence the overlap between the black and gray curves. Because the input stimulus level is smaller on the right ear, the right-ear compressor applies more gain (re linear) to the stimulus than the left-ear compressor does. As a result, $\mathrm{ILD_o} < \mathrm{ILD_i}$. With the MOC strategy (**B**), the output level is larger for the left than for the right ear processor. Therefore, the left ear output inhibits the right-ear compressor more than the other way around. This would turn the right-ear compressor more ‘linear’ with minimal or no change in the left-ear compressor. As a result, $\mathrm{ILD_o}$ would be larger with the MOC than with the STD strategy. Note that in this example, the compression functions were calculated using Eq. (1), and that the input and output levels are in logarithmic scales in dB FS, where 0 dB FS corresponds to a peak amplitude at 1, which itself corresponds to an electrical current at maximum comfortable loudness (MCL). RE: right ear; LE: left ear.

Table 1
Participants' data. Participants whose IDs start with ME and SA were tested in North Carolina and Salamanca, respectively. F: female; M: male; Un: unknown; He: hereditary; Ge: genetic; Mg: meningitis; Ab: antibiotics. L: left; R: right.; pps: pulses per second; MCL: maximum comfortable loudness, B: behavioral.

ID	Sex	Age (years)	Etiology	Time of implant use (months)		Num. Electrodes used for testing		Pulse rate per channel (pps)		Better ear	THR (% MCL)		Vol (%)	
				L	R	L	R	L	R		L	R	L	R
ME115	M	81	Un/He	47	47	9	9	1587.3	1587.3	R	0	0	100	100
ME131	M	54	Un/He	30	32	11	11	1578.9	1823.7	L	0	0	100	100
ME132	M	43	Un	62	62	9	9	1587.3	1587.3	R	B	B	92	100
SA004	F	35	Ge	22	13	11	11	1550	1567	R	10	10	125	120
SA005	M	44	Mg	119	103	11	11	1600	1504	R	0	0	110	100
SA008	M	16	Un	13	129	10	10	1818	1020	R	10	5	130	100
SA009	M	15	Ge	105	148	10	10	1818	1538	R	0	10	125	130
SA010	M	16	Un	140	172	10	10	1695	1099	R	10	0	130	130
SA011	F	44	Un/Ab	22	135	10	10	1754	1734	L	5	5	110	120
SA014	M	48	Mg	175	190	9	9	1846	1143	L	5	5	100	120
SA015	F	35	Mg	147	19	11	11	1405	1653	L	5	5	110	110
SA016	F	74	Un/He	150	119	10	10	1493	1478	L	10	10	110	110

University of Salamanca (Salamanca, Spain). Testing procedures were approved by the Western Institutional Review Board (Puyallup, WA) and by the Ethics Review Board of the University of Salamanca.

2.2. *Stimuli*

Stimuli consisted of Gaussian noise bursts generated digitally (using Matlab's randn function) and bandpass filtered between 125 and 6000 Hz with a fourth-order (North Carolina) or first-order (Salamanca) Butterworth filter to achieve the desired bandwidth. The noise bursts had a duration of 200 ms and were gated with 20-ms (North Carolina) or 50-ms (Salamanca) raised-cosine onset and offset ramps. A linear gain was applied to the noise bursts to achieve the desired presentation level of -20 dB full scale (FS; dB relative to a peak amplitude at unity). For reference, this level corresponds approximately to 70 dB SPL in MED-EL's clinical CI audio processors. For the North Carolina participants, the stimulus level was randomly roved by up to ±2 dB across stimulus presentations; for the Salamanca participants, the stimulus level remained constant across stimulus presentations. The potential implications of this approach are discussed later.

The level-adjusted noise bursts were preceded and followed by silence periods with a duration of 20 ms, making the stimulus duration equal to 240 ms. To simulate a virtual auditory space, stimuli were filtered with diffused-field equalized head-related transfer functions (HRTFs) for a Knowles Electronics Manikin for Acoustics Research (KEMAR) (Gardner and Martin, 1995). The HRTFs were for speakers located in an anechoic chamber, 1 m away from the center of the KEMAR's head. Stimuli were filtered through HRTFs for 0° elevation and for 11 azimuthal angles from $-75°$ to 75° separated by 15°.

This stimulus choice was intended to facilitate a comparison between the present results (which were for a virtual acoustic setting using non-individualized HRTFs and with experimental processing strategies) with previous reports of the performance of BiCI users tested in the free-field with their own clinical devices (Dorman et al., 2014, 2016; see the Discussion).

2.3. *Processing strategies*

The level-adjusted, HRTF-filtered noise bursts were processed through the STD and MOC processing strategies before they were presented to the BiCI participants via direct stimulation (see below). The STD and MOC strategies have been described in detail elsewhere (Lopez-Poveda, 2015; Lopez-Poveda et al., 2016a; Lopez-Poveda and Eustaquio-Martín, 2018) and only a summary is given here.

STD and MOC processors were identical except for the back-end compression stage. Processors were based on the continuous interleaved sampling (CIS) strategy of Wilson et al. (1991). Each processor included a highpass pre-emphasis filter (first-order Butterworth filter with a 3-dB cutoff frequency of 1.2 kHz); a bank of sixth-order Butterworth bandpass filters whose 3-dB cutoff frequencies followed a modified logarithmic distribution between 100 and 8500 Hz; envelope extraction via full-wave rectification and lowpass filtering (fourth-order Butterworth lowpass filter with a 3-dB cutoff frequency of 400 Hz); a logarithmic compression function (fixed for STD and dynamic for MOC processors; see below); and continuous interleaved sampling of compressed envelopes. The number of filters in the bank was identical to the number of electrodes used for testing in the implant (Table 1), and equal between the left- and right-ear processors. Processors did not include a front-end AGC.

The back-end compression function in all processors was as follows (Boyd, 2006):

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

where x and y are the instantaneous envelope input and output amplitudes[1] 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. For the STD processors, c was fixed at 1000 and was identical at the two ears. This value differed slightly from the value of $c = 500$ used by most of the participants in their clinical devices [the exceptions were the three North Carolina participants (ME115, ME131 and ME132), who were using $c = 1000$ in their clinical processors; the right-ear processor of SA011, 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 $c = 1000$]. For the MOC processors, by contrast, the instantaneous value of c varied dynamically in time depending upon the instantaneous time-weighted *output* amplitude[2] from the corresponding frequency channel in the

---

[1] All signal processing was done in the digital domain with signals having instantaneous linear amplitudes in the range $(-1,+1)$ to avoid clipping.

[2] In previous publications about the MOC strategy (e.g., Lopez-Poveda et al., 2016a, 2016b; 2017; Lopez-Poveda and Eustaquio-Martín, 2018), we used the term "output energy". The term "output amplitude" is, however, more accurate.

contralateral processor: the greater the output amplitude, the smaller the value of c (on-frequency inhibition). Specifically, c varied between approximately 30 and 1000 for contralateral output amplitudes of 0 and −20 dB FS, respectively[3] (see Fig. 2 in Lopez-Poveda et al., 2016a for details).

Inspired by the exponential time-course of activation and deactivation of the MOC reflex (Backus and Guinan, 2006), in the MOC strategy, the instantaneous output amplitude from the contralateral processor, $E(t)$, was calculated as the root mean square (RMS) amplitude of the compressed envelope integrated over a preceding exponentially decaying time window with two time constants (τ_a and τ_b, $\tau_a \leq \tau_b$).

To assess the potential benefit of using binaurally coupled back-end compression for sound source localization (the main aim of the study), all 12 participants were tested with two processing strategies:

1. **STD.** A standard strategy involving two independently functioning processors (one per ear), each with fixed back-end compression.
2. **MOC1.** The binaural MOC strategy with fast time constants $\tau_a = \tau_b = 2$ ms, as originally implemented and tested elsewhere (see Lopez-Poveda et al., 2016a, 2016b).

To compare sound source lateralization performance with various implementations of the MOC strategy designed to reflect more or less realistically the inhibitory characteristics of the natural MOC reflex (aim 2 of the study), seven of the 12 participants (SA008, SA009, SA010, SA011, SA014, SA015 and SA016) were tested with two additional strategies:

3. **MOC2.** A MOC strategy with time constants $\tau_a = 2$ ms and $\tau_b = 300$ ms, thus closer to the (slower) time course of activation of the natural MOCR (Backus and Guinan, 2006).
4. **MOC3.** A bandwidth-normalized, slow MOC strategy (see Lopez-Poveda, 2017). In the MOC1 and MOC2 strategies, the control of compression was identical across frequency channels; that is, for each frequency channel, k, the back-end compression parameter, c_k, was a function of the output amplitude, E_k, and the same function, $c_k = f(E_k)$, was used for all frequency channels. As a result, contralateral inhibition was effectively greater for higher than for lower frequency channels, particularly for broadband signals, for two reasons: (1) because the highpass pre-emphasis filter emphasizes higher frequencies and thus the output amplitude was larger for higher than for lower frequency channels; and (2) because high-frequency channels are broader in frequency and pick up more energy (from broadband signals) than lower frequency channels. While this would be roughly consistent with some psychoacoustical studies that have reported larger effects of the MOC reflex at higher than at lower frequencies (e.g., Bacon and Takahashi, 1992; Carlyon and White, 1992; Aguilar et al., 2015), it would be inconsistent with other studies that have reported larger effects of the MOC reflex at lower than at higher frequencies (Kawase et al., 2003; Lilaonitkul and Guinan, 2009; Aguilar et al., 2013).

   In the MOC3 strategy, the same compression control function was used for all frequency channels. Compression, however, was controlled using the output amplitude from each channel normalized to the channel bandwidth, E'_k:

$$E'_k = E_k \cdot \sqrt{\frac{BW_{ref}}{BW_k}} \quad (2)$$

where BW_k is the channel bandwidth, BW_{ref} is the bandwidth of a reference frequency channel, and all variables (E_k, E'_k, BW_k, and BW_{ref}) are in a linear scale. Unless otherwise stated, we chose to make BW_{ref} approximately equal to the bandwidth of the median channel (the actual normalization channel was #7, #6, #5 and #5 for participants with 12, 11, 10 and 9 active channels, respectively.). This produced effectively greater inhibition in the lower than in the higher frequency channels, as shown later and in Fig. 8 of Lopez-Poveda and Eustaquio-Martín (2018). The exception was SA008, who had 10 active electrodes and for whom the normalization channel was #7.

We note that the slow time constant of contralateral inhibition in the MOC2 and MOC3 strategies was longer ($\tau_b = 300$ ms) than the stimulus duration (200 ms). The potential implications of this choice are described below. Further details about all four strategies can be found in Lopez-Poveda and Eustaquio-Martín (2018).

2.4. *Fitting and loudness level balance*

Before any testing, 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 percent of MCL values (Boyd, 2006), according to each participant's preference (Table 1). Post-processor volume controls were used to ensure that sounds at the two ears were perceived as comfortable and equally loud, and that sentences filtered with the HRTF for 0° elevation and 0° azimuth were perceived in the center of the head. The volume setting in each processor adjusted the electrical current at MCL without affecting current threshold, such that a volume at 100% meant that the maximum output current remained at MCL and a volume greater than 100% resulted in a linear scaling-up of the programmed electrical current for MCL (an approach depicted as “Innsbruck” or IBK in Fig. 2 in Boyd, 2006). The volume scaling affected all electrodes equally. Note that a volume setting above 100% was required for some subjects to achieve appropriate loudness levels (Table 1). Thresholds, MCL levels and processor volumes were set with the STD strategy and remained identical for the MOC strategies to ensure that the contralateral inhibition in the MOC strategies produced the corresponding reductions in stimulation amplitudes (i.e., reduced loudness or audibility) relative to the STD condition that would be expected from the natural MOC reflex (e.g., Aguilar et al., 2015). In other words, thresholds, MCL and volumes did not affect MOC processing.

2.5. *Procedure*

For each processing strategy, participants were presented with eight noise tokens for each one of the 11 azimuthal angles (88 noise tokens in total). The 88 noise tokens were presented in random order. During the presentation of the stimuli, participants sat in front of a computer screen that displayed a top view of a human head with an array of speakers in front of the head (Fig. 2). For each stimulus presentation, the subject was instructed to judge the azimuthal position of the sound source by clicking on the corresponding speaker in the computer screen. The click of a response triggered the processing of a freshly generated noise stimulus

---

[3] Note that back-end compression (Eq. (1)) effectively amplifies signal amplitudes. Hence, for noise bursts at −20 dB FS, the levels of the compressed envelopes at the output of the STD strategy can be larger than −20 dB FS. For example, Fig. 4 shows that for a noise sound source at −60° azimuth with a level at −20 dB FS, the level at the output of the STD varies between about −15 and 0 dB FS depending on the ear and the frequency channel.

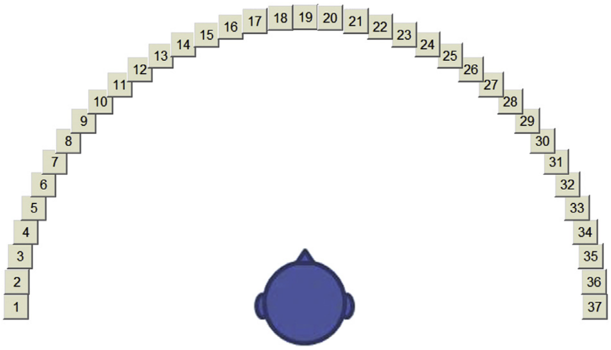

Fig. 2. The response window used in Salamanca.

through the corresponding strategy, and the presentation of the processed stimulus to the participant. In North Carolina, the response screen displayed 11 speakers spaced every 15° over an azimuth range from −75° to 75°. In other words, the range of possible responses was equal to the range of actual azimuth locations. In Salamanca, the screen displayed an array of 37 speakers spaced every 5° over an azimuth range from −90° to 90° (Fig. 2), even though stimuli were presented at azimuths from −75° to 75° every 15°. The latter approach was used to increase the angle error at chance performance (see below).

In North Carolina, tests were conducted in two blocks of four measurements per angle (i.e., two blocks of 44 presentations) per strategy. Participants ME115 and ME132 performed the tests in the order STD, STD, MOC1, and MOC1, while ME131 performed the tests in the order MOC1, STD, STD, and MOC1. In Salamanca, tests were conducted in one block of eight measurements per angle per strategy (88 presentations in total) and the different processing strategies were tested in random order. Additional precautions were taken to minimize potential learning effects that might have biased scores across strategies. First, participants were encouraged to train themselves on the task by clicking on any of 37 speakers evenly spaced every 5° over an azimuth range from −90° to 90° and listening to the corresponding stimulus (that is, during training, participants could hear stimuli at all those azimuthal locations while for testing, stimuli were presented at a subset of locations). Training was provided independently for each processing strategy and for as long as each participant deemed necessary. Second, during the measurements, feedback was not given to participants on the correctness of their responses. Third, participants did not know which processing strategy they were training on or being tested with. Fourth, in Salamanca, the full protocol (four strategies and 88 stimulus presentations per processing strategy) was administered twice for all participants except SA005 and SA008 and the results of the first round were regarded as practice and discarded from further analysis.

Importantly, all tests were ‘double-blind’ such that neither the experimenter nor the participant knew the strategy that was being tested at any time.

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

2.7. Analyses

Response matrices were generated by plotting the reported against the actual azimuth angles. Localization accuracy was quantified using the RMS angle error ($\varepsilon_{\mathrm{RMS}}$), calculated as:

$$\varepsilon_{RMS} = \sqrt{\sum_{i=1}^{N} \frac{(X_i - Y_i)^2}{N}} \tag{3}$$

where X_i and Y_i denote the actual and reported azimuth angles for the i-th stimulus presentation, and N is the total number of presentations ($N = 88$). Localization accuracy was also quantified using the Pearson correlation coefficient between actual and reported azimuth angles, R_{XY}. These two performance metrics are complementary. The correlation coefficient can be advantageous over the RMS angle error when the reported location is systematically to the left or the right of the actual location due to potentially inadequate binaural loudness balance (e.g., see Fig. 3 in Tyler et al., 2006). Conversely, the correlation coefficient is insensitive to potential systematic lateralization bias (i.e., to vertical offsets in the response matrices) that might increase $\varepsilon_{\mathrm{RMS}}$. Both R_{XY} and $\varepsilon_{\mathrm{RMS}}$ are commonly used to quantify accuracy in localization studies (e.g., Majdak et al., 2013; Marmel et al., 2018).

Kolmogorov-Smirnov tests (with Lilliefors correction) were used to test if the distributions of angle error and correlation coefficient were normal. When this happened, parametric repeated-measures analyses of the variance (RMANOVA) and/or paired Student's t tests were used to test for the statistical significance of processing strategy on angle error or correlation scores. When the distributions were not normal, then Friedman and Wilcoxon signed-rank tests were conducted instead. An effect was regarded as statistically significant when the null hypotheses could be rejected with 95% confidence ($p \leq 0.05$). For tests involving multiple groups or variables, post hoc pairwise comparisons were conducted using Bonferroni corrections of the p value for multiple comparisons. All statistical tests were conducted in IBM SPSS Statistics v23.

3. Results

In this section, we first analyze the level cues provided by the different processing strategies and their time course. We then compare localization scores for the (originally proposed) MOC1 strategy with those for the STD strategy (aim 1 of the study). Lastly, we compare localization scores for the various implementations of the MOC strategy (MOC1, MOC2 and MOC3) with those for the STD strategy (aim 2 of the study).

3.1. Level cues provided by the STD and MOC strategies

Fig. 3 shows output envelopes for a sound source located at −60° azimuth from STD, MOC1, MOC2 and MOC3 processors with 10 frequency channels, the typical number of channels for the present participants (Table 1). For conciseness, output envelopes are shown for three channels only: channel #3 (bottom row), #5 (middle row) and #10 (top row), with center frequencies of 501, 1159 and 7230 Hz, respectively. The left and middle columns illustrate output amplitudes for the left ear and right ear, respectively, and the right-most column illustrates the difference in output amplitude between the left and the right ears (note that this

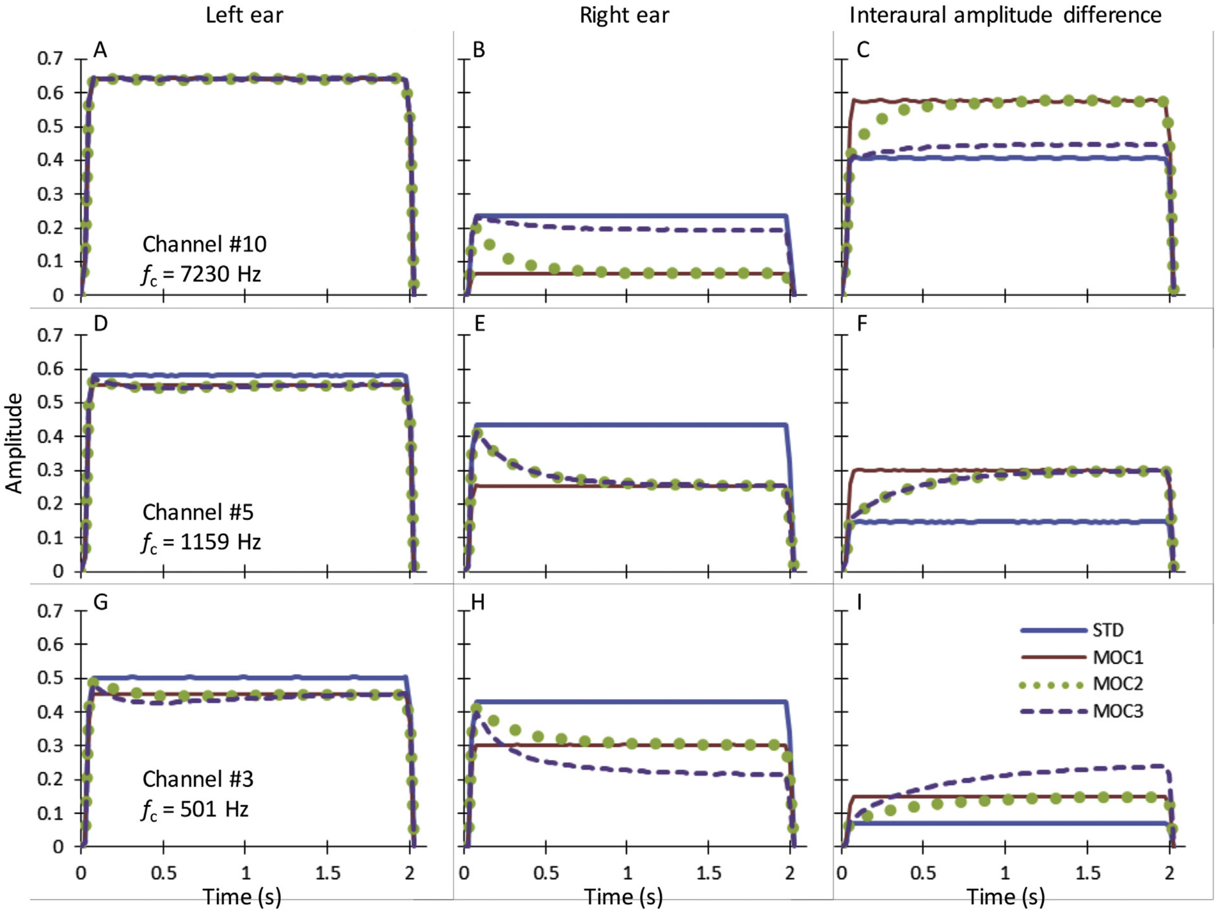

Fig. 3. Example output signals (compressed envelopes) and interaural amplitude difference for STD, MOC1, MOC2 and MOC3 processors with 10 frequency channels, and for a sound source located at −60° azimuth. The MOC3 strategy was implemented with $BW_{\mathrm{ref}} = BW_{\#5}$. The stimulus was a ten-tone complex (2 s in duration with 50-ms onset and offset ramps), with all tones having identical input level. The overall stimulus level was −20 dB FS. Each row shows signals for a different frequency channel with center frequencies (f_c) of 501 Hz (channel #3, bottom row), 1159 Hz (channel #5, middle row), and 7230 Hz (channel #10, top row). **Left column.** Amplitude at the output of the left ear processor. **Middle column.** Amplitude at the output of the right ear processor. **Right column.** Difference in output amplitude between the left and the right ear. Each panel illustrates four traces (one per processing strategy), as indicated by inset in panel I. See main text for details.

is different from the output ILD, which is discussed below). To better illustrate the effects of contralateral inhibition in the MOC strategies, the stimulus consisted of ten pure tones equal in amplitude and whose frequencies were approximately at the center of the processors' frequency channels. The stimulus was long enough (its duration was 2 s with 50-ms cosine-squared onset and offset ramps) to reveal the full inhibitory effects of the slower MOC strategies, and its overall level was set at −20 dB FS, thus equal to the level of the noise bursts used in the localization experiments. To facilitate the visualization of the different traces using different line styles, the output signals were first smoothed (using Matlab's smooth function) and then downsampled from 20 kHz to 40 Hz.

The figure illustrates the following:

1. For any given processing strategy and frequency channel, the output amplitude was greater for the left ear (left column in Fig. 3) than for the right ear (middle column in Fig. 3). This is because the sound source was located on the left side of the head (at −60° azimuth) and the HRTF introduced a head shadow ILD. The interaural amplitude difference (right panels in Fig. 3) was larger for channel #10 than for channels #3 or #5 because channel #10 was higher in frequency and the head shadow ILD is greater at higher than at lower frequencies (e.g., Blauert, 1997; Lopez-Poveda, 1996).
2. In the right ear (the shadowed ear in this example; middle column in Fig. 3), the amplitude was always greater or equal for the STD strategy than for any of the MOC strategies. This is because in the MOC strategies, the ear with the largest output amplitude inhibits the ear with the smallest output amplitude more than the other way around (see Methods). Because the output amplitude in this example was greater for the left ear than for the right ear, the left ear inhibited the right ear more than the other way around, which reduced the output amplitude more in the right than in the left ear.
3. Contralateral inhibition was faster for the MOC1 than for the MOC2 or MOC3 strategies. For the MOC2 or MOC3 strategies, it took approximately 1 s for the output amplitude in the right ear to achieve its asymptotic value. The different time course between the MOC strategies was related to using a faster time constant of contralateral inhibition in the MOC1 than in the MOC2 or MOC3 strategies (see Methods).
4. The interaural amplitude difference (right column in Fig. 3) was equal or greater for any MOC strategy than for the STD strategy. This is because contralateral inhibition in the MOC strategies reduced the output amplitude in the right ear.
5. In the high frequency channel #10, the asymptotic interaural amplitude difference was greater with the MOC1 and MOC2 strategies than with the MOC3 strategy (Fig. 3C). The opposite

was true for channel #3 (Fig. 3I). For channel #5 (Fig. 3F), all three strategies produced equal amount of contralateral inhibition, thus an equal interaural amplitude difference in the asymptote. The different effect of MOC processing on the interaural amplitude difference was related to using (or not) bandwidth normalization (see Methods).

Fig. 4 illustrates the RMS output level (computed over the whole stimulus duration and expressed in dB FS) and the ILD (in dB) for each frequency channel and for various processors with 10 frequency channels. In this example, stimuli were identical (200-ms, wideband noise bursts) as those used in the experiments and the sound source was located at −60° azimuth. The top and middle panels illustrate output levels for the left and right ears, respectively; the bottom panel illustrates the ILD calculated as $20 \times \log_{10}(O_{LE}/O_{RE})$, with O_{LE} and O_{RE} denoting the RMS output amplitudes (in linear units) at the left and right ears, respectively. To illustrate the effect of compression, Fig. 4 also shows the output levels and ILDs for a 'linear' STD strategy without back-end compression [achieved by setting $c = 1\mathrm{e}{-}10$ in Eq. (1)]. The peak and valleys of the STD-LIN trace reflect the stimulus spectrum combined with the spectral shape of the HRTF, the highpass pre-emphasis filter, and the filter bank. The figure shows that for all strategies and channels, the output level was greater or equal for the left ear than for the right ear. This is because the sound source was located on the left side of the head and the HRTF introduced a head shadow ILD. In addition, the output level was greater for any strategy than for STD-LIN because all four strategies (STD, MOC1, MOC2 and MOC3) applied back-end compression that amplified the lower input levels more than the high input levels. Compression, however, reduced the spectral contrast at each ear as well as the ILD (Fig. 4C). We note that of the four test strategies, the MOC1 strategy provided spectral contrast and ILDs that were most similar to the values that would be available without the 'detrimental' effects of compression (depicted as STD-LIN in Fig. 4). We also note that the MOC2 and MOC3 strategies provided similar output levels, ILDs and spectral contrast as the STD strategy did because the stimulus was shorter (200 ms) than the time required for full activation of contralateral inhibition in the MOC2 and MOC3 strategies (Fig. 3).

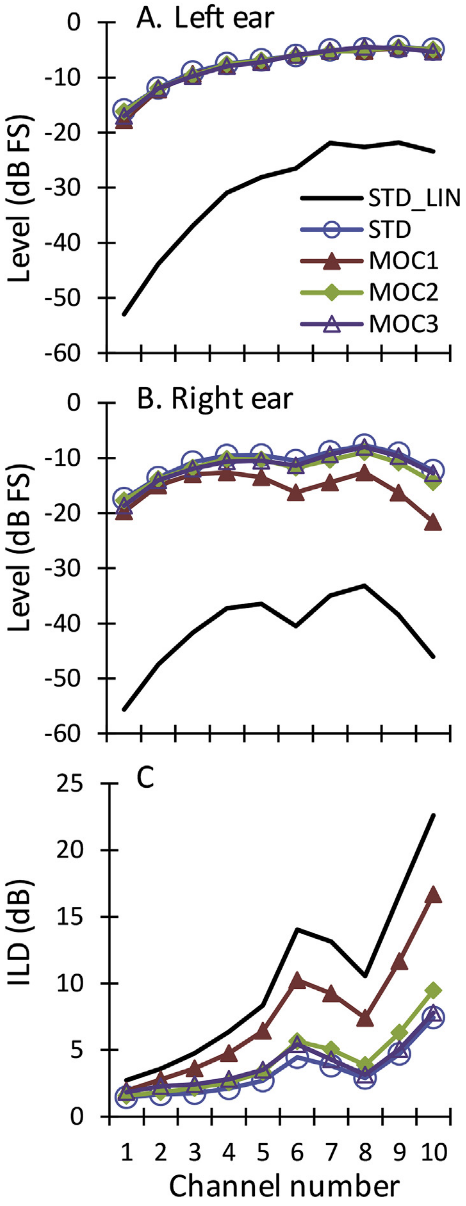

Fig. 4. Output level (in dB FS) at the left ear (**A**), right ear (**B**), and ILD (**C**) as a function of channel number for STD, MOC1, MOC2, and MOC3 sound processors with 10 frequency channels. Also shown are the levels and ILD for a linear STD processor with minimal back-end compression (STD_LIN). The stimulus was a 200-ms wideband noise burst identical to those used for testing and the source was located at −60° azimuth. The MOC3 strategy was implemented with $BW_{\mathrm{ref}} = BW_{\#5}$.

Fig. 5 illustrates the overall output levels at each ear as well as the ILD as a function of azimuth angle for STD, MOC1, MOC2, and MOC3 strategies with 10 frequency channels. Stimuli were identical (200-ms, wideband noise bursts) as those used in the experiments and were vocoded (using noise carriers) through the corresponding processing strategy. The vocoder has been described elsewhere (Lopez-Poveda and Eustaquio-Martín, 2018). For reference, the figure also shows the stimulus levels at the input of the processors (i.e., after HRTF filtering), depicted as HRTF. Because the MOC strategies are sensitive to stimulus levels (see Lopez-Poveda, 2015), results are shown for stimulus levels of −40, −30, −20 and −10 dB FS, as indicated at the top of each column. (Note that the present experiments with BiCI users were conducted with stimuli around −20 dB FS).

Fig. 5 illustrates the following:

1. For all strategies, the overall output levels at each ear increased gradually with increasing stimulus level (i.e., levels increased from the left-most to the right-most panels in the top and middle rows of Fig. 5). However, the difference between the input (HRTF) and the output levels decreased with increasing stimulus level (i.e., the length of the vertical arrows in the top and middle rows of Fig. 5 decreased from left to right). This is because back-end compression (Eq. (1)) amplified the lower input levels more than the higher input levels.
2. At each ear, the MOC1 strategy produced the steepest level-azimuth functions, the more similar in slope to the corresponding HRTF functions, and the more constant in slope across the range of stimulus levels tested (i.e., the dashed lines and the filled triangles in the top and middle rows had identical or very similar slopes from left to right). The STD, MOC2 and MOC3 strategies, by contrast, produced level-azimuth functions that became gradually shallower as the stimulus level increased (their slope decreased from left to right in Fig. 5). In other words, the MOC1 strategy preserved to a larger extent the monaural HRTF level localization cues across the range of stimulus levels tested. For the STD strategy, these monaural level cues decreased gradually with increasing sound level because compression enhanced the lower input levels in the shadowed ear more than the higher input levels in the ear closer to the

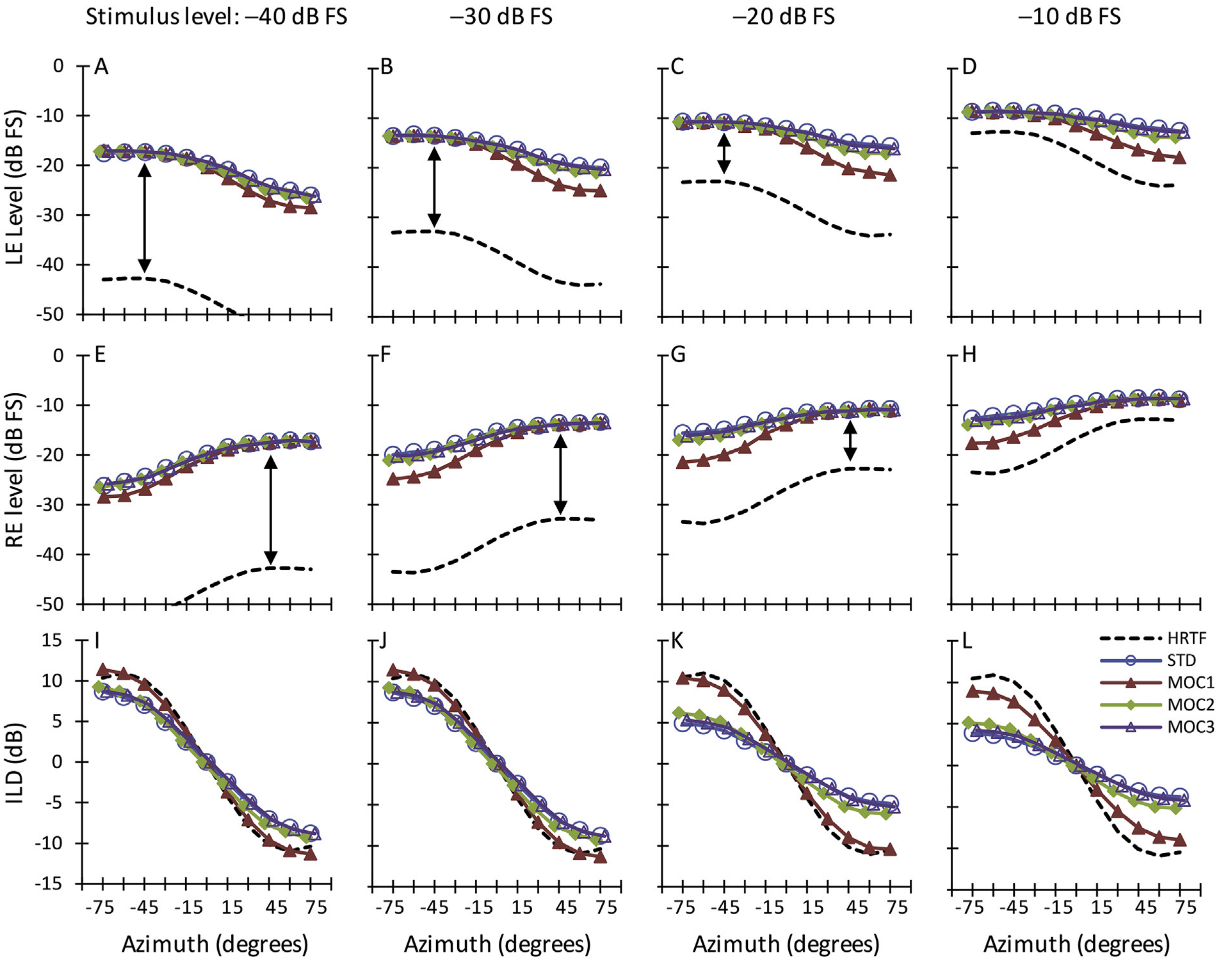

Fig. 5. Overall output level at the left ear (LE, **top row**), right ear (RE, **middle row**), and ILD (**bottom row**) as a function of azimuth location for noise-vocoded stimuli with STD, MOC1, MOC2, and MOC3 sound processors with 10 frequency channels. The MOC3 strategy was implemented with $BW_{\mathrm{ref}} = BW_{\#5}$. Also shown are the amplitudes at the input of the processors and the corresponding ILDs, depicted as HRTF. Each column shows results for a different stimulus level, from −40 to −10 dB FS, as indicated at the top of the column.

sound source. While the contralateral inhibition of compression used in all MOC strategies can theoretically preserve those monaural cues, only the MOC1 strategy preserved those cues because only for this strategy was contralateral inhibition maximally active over the virtually full stimulus duration (i.e., as noted earlier and in Fig. 3, the stimulus duration was shorter than the time required for full activation of contralateral inhibition in the MOC2 or MOC3 strategies).

3. The MOC1 strategy produced the largest ILD, the closest to the HRTF ILD, and the more constant across the range of stimulus levels tested. By contrast, the ILD was comparable for the MOC2, MOC3 and STD strategies, was smaller than the HRTF ILD, and nearly halved as the stimulus level increased from −40 to −10 dB FS. The ILD was largest for the MOC1 strategy because only for the MOC1 strategy was contralateral inhibition maximally active over virtually the full stimulus duration (see Fig. 3).

3.2. Localization with the MOC1 and STD strategies

Neither the angle error nor the correlation coefficient for each individual participant from the smaller North Carolina group (N = 3) were outside the mean plus or minus two standard deviations interval for the more numerous Salamanca group (N = 9). Furthermore, the North Carolina and Salamanca groups were not significantly different in mean angle error or correlation coefficient with either the STD or the MOC1 strategies (two-tailed Student *t* tests for unequal sample sizes with unequal variances produced *p* values of 0.45 for the difference in mean angle error between the two groups with the STD strategy, 0.13 for the difference in mean angle error with the MOC1 strategy; 0.56 for the difference in mean correlation coefficient with the STD strategy; and 0.95 for the difference in mean correlation coefficient with the MOC1 strategy). This justified analyzing the data for the Salamanca and North Carolina participants jointly.

Fig. 6 shows example response matrices for two example participants: the 'best' overall performer with the smallest angle errors (SA004, top panels) and a typical performer with angle error scores close to the group mean scores (SA014, bottom panels).

Fig. 7 illustrates individual and group mean localization angle error scores ($\varepsilon_{\mathrm{RMS}}$, Eq. (3)) for the MOC1 and STD strategies. Chance performance for the North Carolina and Salamanca setups (calculated by assessing random localization performance) was approximately 64° and 70°, respectively. All participants performed better than chance. For eight of the 12 participants, the angle error was smaller for the MOC1 than for the STD strategy. For participants SA005, SA009, SA011, and SA015, the angle error was comparable for the two strategies. Kolmogorov-Smirnov tests (with Lilliefors correction) revealed that angle error scores for the STD and MOC1 strategies each conformed to a normal distribution ($p > 0.200$), thus

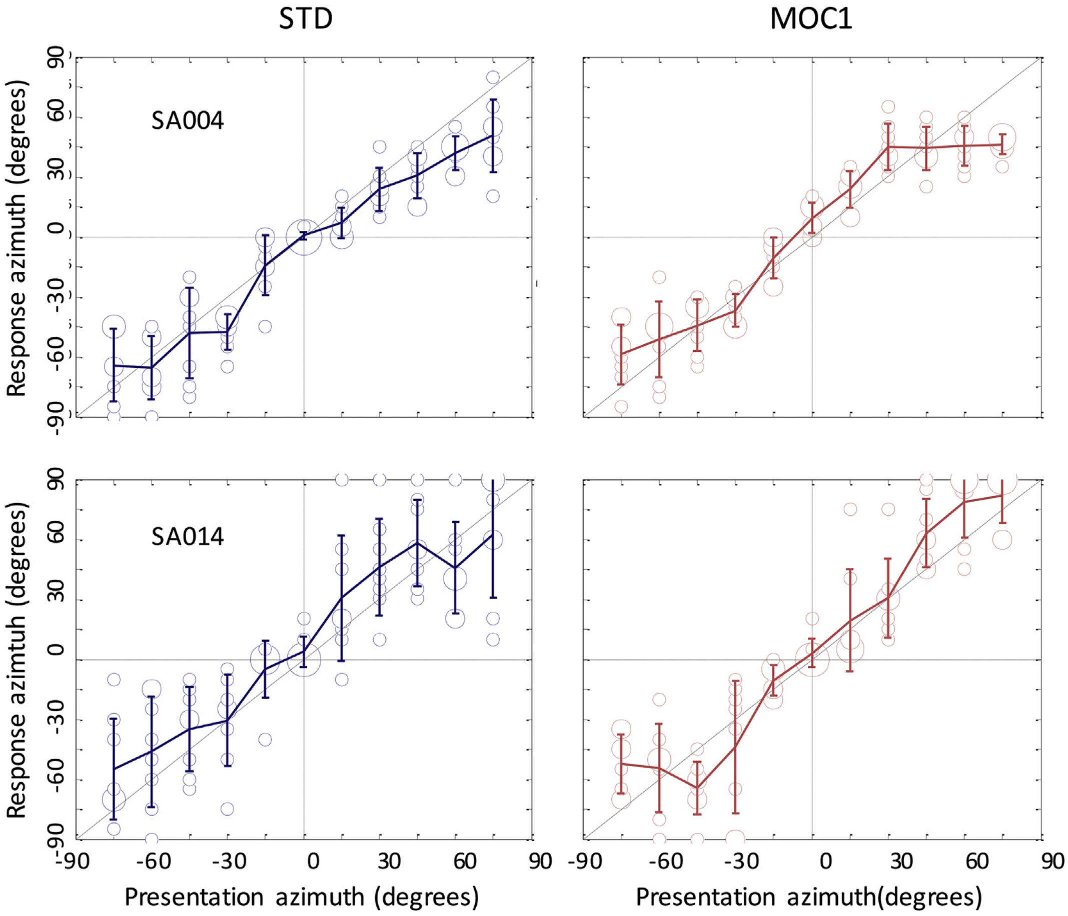

Fig. 6. Example localization matrices for two BiCI users: SA004 (top) and SA014 (bottom). A pair of matrices is shown for each participant: one for the STD strategy (left) and one for the MOC1 strategy (right). Within each matrix, the reported azimuth is shown as a function of presented azimuth angle. Eight stimuli were presented for each azimuth angle. The size of each point is proportional to the number of responses at the corresponding angle. Lines show the mean reported angle for every actual angle and error bars illustrate one standard deviation.

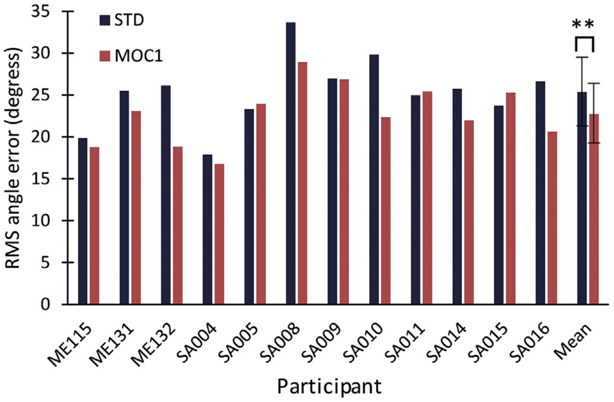

Fig. 7. Angle error for the MOC1 and STD strategies, as indicated by the inset. Results are shown for each individual participant and the mean across participants. Lower values indicate better performance. Error bars for the mean scores depict one standard deviation. **: $p \leq 0.01$.

it was justified to use parametric statistical tests to compare angle error score with the MOC1 and STD strategies. The group mean angle error score was smaller for the MOC1 (mean ± s.d. = 22.7°±3.6°) than for the STD strategy (25.3°±4.1°) and the difference was statistically significant (two-tailed, paired Student's *t*-test, $p = 0.0015$, $N = 12$).

Fig. 8 shows the correlation coefficient (R_{XY}) between the actual and reported azimuth for the MOC1 and the STD strategies for each individual participant and the mean across participants. Nine participants showed higher (better) correlation coefficients with the MOC1 than with the STD strategy and three participants (SA004, SA005 and SA008) showed similar correlation coefficients with the two strategies. Kolmogorov-Smirnov tests (with Lilliefors correction) revealed that the correlation coefficients for the STD and MOC1 strategies each conformed to a normal distribution ($p > 0.200$). The group mean correlation coefficient was higher (better) with the MOC1 (mean ± s.d. = 0.92 ± 0.024) than with the STD strategy (0.89 ± 0.037) and the difference was statistically significant (two-tailed, paired Student's *t*-test, $p = 0.005$, $N = 12$).

Fig. 9A allows a comparison of group mean angle error scores for the MOC1 and the STD strategy for each azimuth location. The two strategies produced similar errors (within ±2°) for azimuths at or near ±30° (Fig. 9B). The MOC1 strategy, however, tended to improve lateralization for virtually every other azimuth, particularly for sources near the midline (i.e., for azimuths between $-15°$ and $+15°$) and on the far sides (i.e., for azimuths $\geq +60°$ or $\leq -60°$). A RMANOVA revealed a statistically significant effect of processing strategy [$F(1,11) = 10.52$, $p = 0.008$]. However, neither the effect of angle [$F(10,110) = 1.37$, $p = 0.220$] nor the interaction between processing strategy and angle were statistically significant [$F(10,110) = 0.85$, $p = 0.581$].

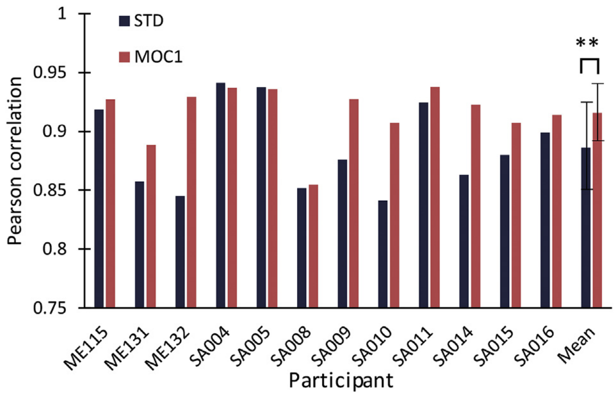

Fig. 8. Correlation between presentation and response azimuth for the MOC1 and STD strategies, as indicated by the inset. Results are shown for each individual participant and the mean. Higher values indicate better performance. Note that the ordinate scale starts at 0.75 rather than zero to better show small differences. Error bars for the mean scores illustrate one standard deviation. **: $p \leq 0.01$.

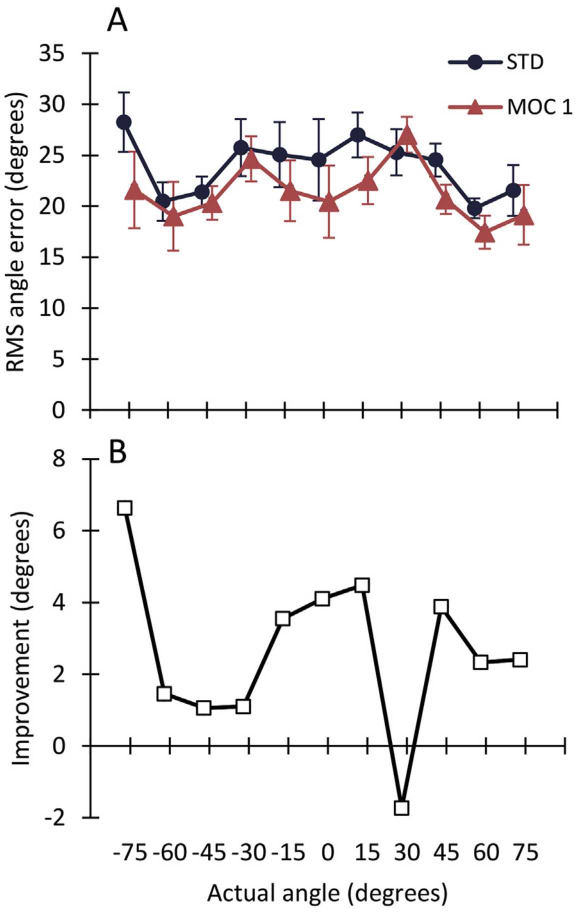

Fig. 9. **A.** Mean RMS error angle for each azimuth location and for the STD and MOC1 strategies. Error bars illustrate one standard error of the mean ($N = 12$). Data points have been slightly displaced horizontally to reduce overlap. **B.** Mean localization improvement with the MOC1 strategy, calculated as the mean angle error for the STD strategy minus the mean angle error for the MOC1 strategy, all in degrees.

3.3. Localization with various implementations of the MOC strategy

Seven of the 12 participants (SA008, SA009, SA010, SA011, SA014, SA015, and SA016) were tested with the STD, MOC1, MOC2 and MOC3 strategies. Fig. 10 allows a comparison of localization performance with the different strategies. The trends were different for different participants.

The worst performance occurred for participant SA008 with the MOC3 strategy (angle error = 35°, correlation = 0.75). This was probably due to two factors. First, SA008 was the worst performer overall, regardless of the strategy. Second, in implementing the MOC3 strategy, BW_{ref} was made equal to $BW_{\#7}$ for participant SA008, while it was approximately equal to the *BW* of the median channel for all other participants (see Methods). Participant SA008 had 10 active electrodes and so normalizing to $BW_{\#7}$ probably caused excessive inhibition that compromised audibility, which may have degraded his performance. For these reasons, MOC3 scores for participant SA008 were omitted from the mean values in Fig. 10 and from the following statistical analyses.

Kolmogorov-Smirnov tests (with Lilliefors correction) showed that angle error scores conformed to a normal distribution for all four strategies (mean ± s.d. for STD: 27.4°±3.1°, $p = 0.150$; MOC1: 24.5°±2.7°, $p > 0.200$; MOC2: 27.2°±3.8°, $p = 0.135$; MOC3: 24.6°±2.5°, $p > 0.200$), thus it was justified to use a RMANOVA to test for the effect strategy on angle error score. The RMANOVA test revealed no significant effect of processing strategy on angle error [$F(3,15) = 1.49$, $p = 0.26$].

Kolmogorov-Smirnov tests (with Lilliefors correction) revealed that the correlation coefficient conformed to a normal distribution for the STD (0.88 ± 0.026, $p > 0.200$), MOC2 (0.88 ± 0.036, $p > 0.200$), and MOC3 (0.89 ± 0.024, $p > 0.200$) strategies, but not for the MOC1 strategy (0.91 ± 0.025, $p = 0.011$). A Friedman test revealed a statistically significant difference in correlation between actual and reported azimuth depending on the strategy [$\chi^2(3) = 9.343$, $p = 0.025$]. Post-hoc pairwise analysis with Wilcoxon signed-rank revealed a trend for better (higher) correlation with the MOC1 than with any other processing strategy (STD vs. MOC1: $Z = -2.521$, $p = 0.012$; STD vs. MOC2: $Z = -0.140$, $p = 0.889$; STD vs. MOC3: $Z = -0.507$, $p = 0.612$; MOC1 vs. MOC2: $Z = -2.240$, $p = 0.025$; MOC1 vs. MOC3: $Z = -2.197$, $p = 0.028$; MOC2 vs. MOC3: $Z = -0.338$, $p = 0.735$). However, none of the pairwise comparisons would remain as statistically significant after Bonferroni correction for multiple comparisons [i.e., none of the *p* values was smaller than the corrected critical $p < 0.0083$ (=0.05/6)].

4. Discussion

We have shown that, compared to using two independently functioning CI processors with fixed back-end compression (the STD strategy), the MOC1 strategy (with fast, binaurally coupled dynamic compression) enhances ILD cues (Figs. 4 and 5) and improves the localization of acoustic stimuli in a virtual horizontal plane (Figs. 7 and 8). Alternative implementations of the MOC1 strategy with slower (longer) time constants of integration (MOC2) and with greater inhibition in the lower than in the higher frequency channels (MOC3) can also enhance ILD cues for sufficiently long stimuli (Fig. 3). However, these (more realistic) implementations of the MOC strategy did not enhance the ILDs (Fig. 5) and did not improve the localization of the (short) 200-ms noise bursts used here relative to the STD strategy (Fig. 10).

4.1. Interpretation

BiCI users rely mostly on ILD cues to judge the location of sound sources in the horizontal plane (Dorman et al., 2016; Laback et al., 2004; Litovsky and Gordon, 2016; Seeber and Fastl, 2008). Consistent with this, many aspects of the present results appear to be explained by the ILD versus azimuth functions produced by the tested strategies (Fig. 5K). For example, response matrices tended to flatten from azimuths of ±60° (Fig. 6) possibly because all

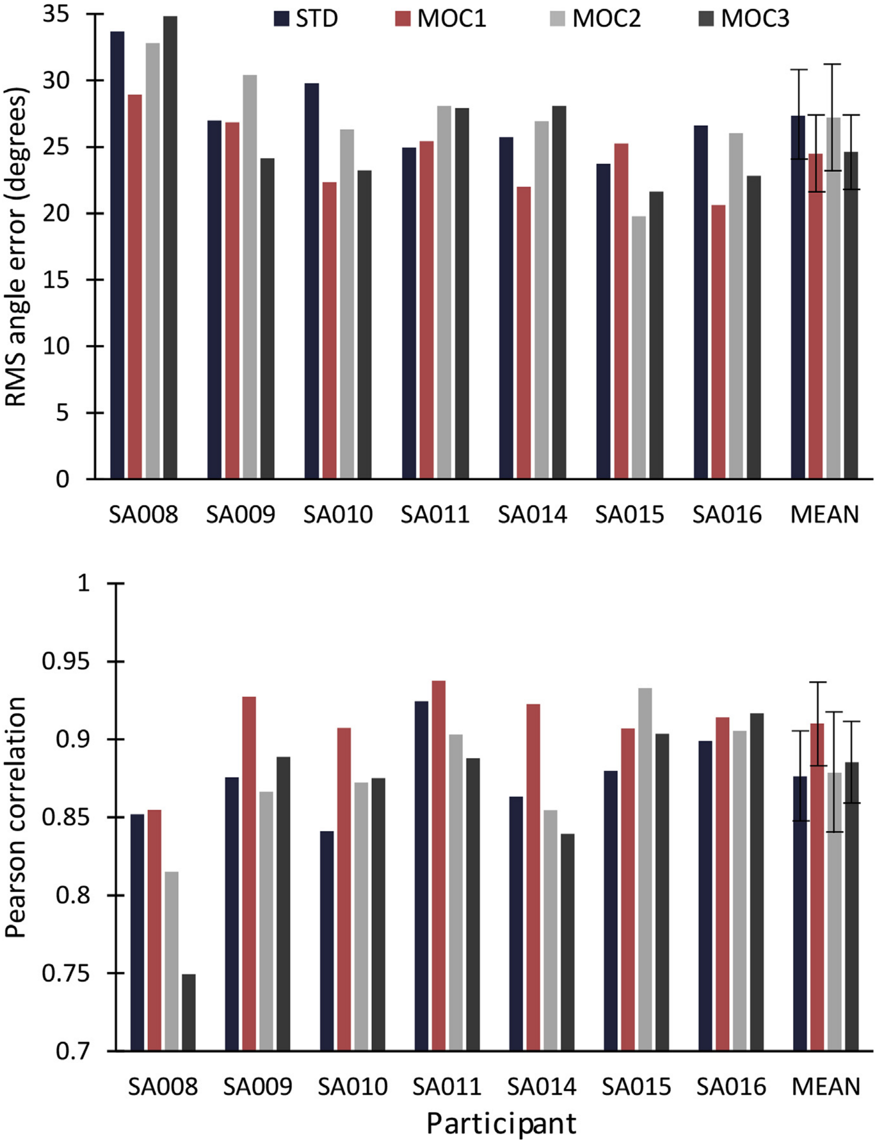

Fig. 10. Angle error (**A**) and Pearson correlation (**B**) between presentation and response angles for the STD, MOC1, MOC2 and MOC3 strategies. Values are shown for individual participants and the mean across participants. Error bars for the mean scores illustrate one standard deviation. MOC3 scores for participant SA008 are not included in the mean scores for the MOC3 strategy because the MOC3 strategy for this participant was implemented with the wrong BW_{ref} (see main text for details). Note that the scale in the ordinate of the bottom panel starts at 0.7 rather than zero to better show smaller differences. Individual scores for the STD and MOC1 strategies are re-plotted from Figs. 7 and 8.

strategies produced roughly constant ILDs for sound sources at and beyond 60°. RMS angle errors tended to be smaller at 60° (Fig. 9A) possibly because stimuli (and response screens) were bounded at 75°/90° and ILDs were approximately constant for azimuths ≥60°, leading listeners to response at 60°.

For the present (200-ms long) stimuli, the MOC1 strategy produced the largest ILDs and the steepest ILD-versus-azimuth function. Furthermore, ILDs for the MOC2, MOC3 and STD strategies were smaller than for the MOC1 strategy, and the corresponding ILD-versus-azimuth functions were shallower (Fig. 5K). This suggests that localization was overall better with the MOC1 strategy because this strategy provided ILDs that were larger than the just-noticeable difference (JND) in ILD for the participants and coded for azimuth less ambiguously than any other strategy did.

The interaction between RMS angle error and azimuth was not statistically significant. Nonetheless, the localization improvements with the MOC1 strategy (re STD) tended to be larger in the frontal region (azimuths between −15° to 15°; Fig. 9), which is also the area with the smaller ILDs (Fig. 5K). This may be because at small angles, the ILDs provided by the STD strategy were smaller than or close to the JND-ILD of the listeners and became discernible with the MOC1 strategy (note that the JND-ILD is smaller at small angles; e.g., Fig. 3 in Yost and Dye, 1988).

4.2. Limitations

In measuring sound localization performance, it is common practice to rove the level of the acoustic stimulus to maximize the chance that localization be based on a ‘true’ interaural level cue rather than on the absolute level at either ear (e.g., Seeber et al., 2004; Majdak et al., 2011). Here, we roved the stimulus level for the three participants tested in North Carolina only but not for the nine participants tested in Salamanca. It is unlikely, however, that conclusions would have been different if we had roved the level for all participants. First, the monaural level-versus-azimuth functions at either ear were shallower than the corresponding ILD-versus-azimuth functions (Fig. 5). For example, for the MOC1 strategy and a stimulus level of −20 dB FS, the level at any ear changed by less than 10 dB over the −60° to 60° azimuth range while the corresponding ILD change was about 20 dB. This held true over a stimulus range from −40 to −10 dB FS. This indicates that the ILD was a more salient and possibly less ambiguous localization cue than the level at any single ear, even with roving of the stimulus level. Second, the trends in the data for the three participants tested with level roving was similar as for the other participants or the mean (e.g., angle errors were smaller, and correlations were greater with the MOC1 than with the STD strategy). However, monaural level cues might have been sufficient for localization if the level change across azimuths exceeded the level JND of the listener, particularly for the MOC1 strategy because it produced the steeper level-versus-azimuth functions (Fig. 5). Therefore, we cannot entirely rule out that the task could be performed to some uncertain extent by monitoring the stimulus level at a single ear.

The stimulus duration (200 ms) was shorter than the time required for a full activation (and deactivation) of contralateral inhibition in the MOC2 or MOC3 strategies (Fig. 3). As a result, the ILDs for the present stimuli were probably smaller than they would have been for longer stimuli. Indeed, vocoder simulations (not shown) revealed that the overall ILD for azimuth angles of ±60° would have been about 3 dB larger for the MOC2 strategy and about 1 dB larger for the MOC3 strategy if the stimulus duration had been 2 s rather than 200 ms (note that the use of longer stimuli would hardly increase the ILDs produced by the MOC1 strategy because contralateral inhibition was very fast in this strategy; i.e., τ_a and τ_b were 2 ms). Therefore, it is possible that the use of longer stimuli might improve localization performance with the MOC2 and MOC3 strategies to some uncertain extent.

Many individual participants showed better localization with the MOC1 than with the STD strategy (Figs. 7 and 8) even though the STD strategy was the most similar to the audio processing strategies worn by the participants in their clinical devices and participants were not given much opportunity to become fully accustomed to MOC processing before testing. Fig. 11 compares angle error scores across the practice session and the data collection session for those participants who had the two sessions. Error scores tended to be smaller in session 2 than in session 1 (i.e., most data points are below the diagonal), suggesting that performance tended to improve with practice. In addition, the vertical offset from the diagonal tended to be larger for participants with larger angle errors in the practice session, suggesting that practice benefited those who performed worse in the first session more than those who performed well. Last, the difference in performance across the two sessions (i.e., the potential effect of practice) tended to be smaller for the STD strategy than for any of the MOC strategies, possibly because the STD strategy provided localization cues most similar to those provided by the participants' own clinical devices. Altogether, this suggests that the potential benefits from the MOC1 strategy (and MOC processing in general) could become larger with practice and/or a sustained use of the MOC strategies.

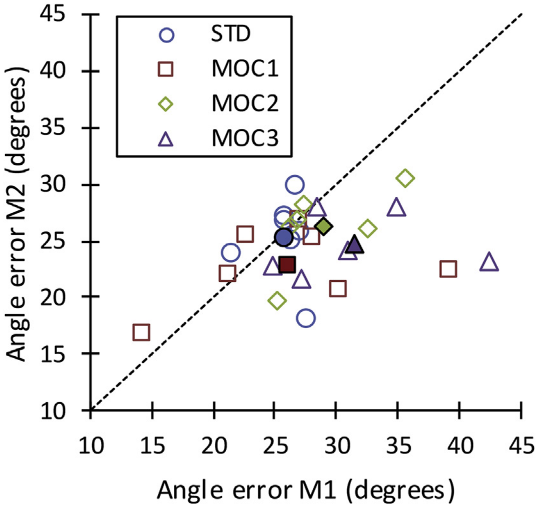

Fig. 11. Comparison of angle error scores for the practice measurement session (M1, abscissa) and the test session (M2, ordinate). Different symbols illustrate results for different processing strategies, as shown by the inset. Open symbols illustrate results for six (MOC2 and MOC3 strategies) or seven (STD and MOC1 strategies) individual participants and filled symbols illustrate group mean results.

4.3. Comparison with related studies

The present tests were in simulated free-field conditions and the processing strategy used as the reference (the STD strategy) may have differed from the processing employed by the participants in their clinical devices. One might wonder (1) to what extent are the present results representative of lateralization in the free field? And (2) to what extent may the present findings generalize to clinical audio devices?

For listeners with normal hearing, the use of non-individualized HRTFs degrades the spectral details responsible for determining sound source elevation (e.g., Marmel et al., 2018) but to a large extent preserves the interaural difference cues responsible for determining the location of a sound source in the horizontal plane (Wenzel et al., 1993). Francart et al. (2011) reported that for listeners wearing a CI in one ear and a hearing aid in the other ear, the mean angle error for lateralization in a virtual sound field (28.4°) did not differ from that in the real sound field (31.5°). The angle errors obtained here with the reference, STD strategy (18° to 33°, mean = 25.3°, Fig. 7) were within the range of values reported in the literature for BiCI users tested with their clinical audio processors and for sound sources in the free field spanning a (broad) azimuth range similar to the one used here [e.g., the mean angle error in the free field was 24.5° in Nopp et al. (2004); 24.0° in Verschuur et al. (2005); 24.1° for noise and 21.5° for speech signals in Grantham et al. (2007); 20.4° for a wideband signal, 19.6° for a highpass signal, and 43.4° for a lowpass signal in Dorman et al. (2014); or 29.0° in Dorman et al. (2016)]. While some studies have reported smaller angle errors [e.g., 10° in van Hoesel and Tyler (2003)] or BiCI users performing close to normal (e.g., Seeber et al., 2004; Seeber and Fastl, 2008), this was probably due to using a narrower azimuth range over which the ILD-vs-angle is monotonic. For example, van Hoesel and Tyler (2003) used eight loudspeakers spaced at 15.5° and spanning 108° in front of the participant, and Seeber et al. (2004) and Seeber and Fastl (2008) used 11 speakers spaced at 10° from −50° to 50°. Therefore, altogether it is unlikely that the use of non-individualized HRTFs affected localization significantly. Even if it did, the effects of using non-individualized HRTFs would have been comparable across processing strategies. Altogether, this supports the conclusion that (1) the present results for the STD strategy are likely representative of the results that would be obtained with clinical devices in the free field; and (2) that it would not be unreasonable to generalize the reported effects of processing strategy to free-field tests.

We note, however, that the present tests were conducted without a front-end AGC. This differs from most clinical audio processors, which include a front-end broadband AGC compression stage (Zeng, 2004). In addition, we balanced the volume at the two ears to ensure that sentences at the two ears were perceived equally loud. This differs from typical clinical practice, where the output volume of each processor is set independently (Ching et al., 2007; Tyler et al., 2006). In addition, the HRTFs employed here were for a KEMAR (thus nonindividualized) and were recorded with microphones placed at the eardrum position in a minimally reverberant (anechoic) room (Gardner and Martin, 1995). Therefore, the present HRTFs almost certainly provided different localization cues (input ILDs were possibly larger) than the participants were used to with their clinical audio processors in realistic, reverberant listening conditions. While binaural loudness balancing would seem appropriate in clinical practice and excluding AGC seems reasonable for isolating the effects of back-end compression on localization, participants may have adapted to different ILD-to-angle functions than they were used to with their devices in daily life.

BiCI users lateralized more accurately with the MOC1 than with the STD strategy (mean angle error was 22.7° versus 25.3°, respectively). The MOC strategies were designed to reinstate the contralateral, dynamic control of compression mediated by the natural contralateral MOC reflex, which is absent for BiCI users (Lopez-Poveda et al., 2016b; Lopez-Poveda, 2018). If successful, one would expect the performance of BiCI users with the MOC strategies to be closer to the performance of listeners with normal hearing in the same task. The comparison remains to be done. It is unlikely, however, BiCI users would show normal localization accuracy with the MOC1 strategy in realistic free-field settings. In natural listening environments, normal-hearing listeners have access to individualized ITD, ILD and spectral cues that would still be absent to BiCI users with the MOC strategy. Dorman et al. (2016) reported that in a free-field localization task with stimuli identical to the stimuli employed here and with a similar speaker arrangement, mean angle error scores were significantly greater (worse) for BiCI users than for young, normal-hearing listeners (29° versus 6°) and even the 'best' BiCI users had error scores above the 95th percentile of scores for young, normal-hearing listeners. This suggests that the mean angle error improvement of 2.6° provided by the 'best' MOC1 strategy would be insufficient to bring the performance of BiCI users equal to that of normal-hearing listeners, even if BiCI users were given sufficient practice on the MOC1 strategy.

4.4. Comparison with other binaural processing strategies

Several ILD-enhancement methods have been recently proposed. For example, Moore et al. (2016) proposed to enhance ILDs at low frequencies (≤1500 Hz; the frequency range where HRTF ILDs are smaller and ITDs are greater) by 'mapping' inter-aural phase differences into ILDs. The method was expected to create a (correct) perception of sound source location, even if the listener is insensitive to the corresponding ITD, as might be the case for typical BiCI users. The method was tested on bilateral hearing-aid users using simulated hearing aids. The algorithm did not improve the localization of noise sources but improved the localization of speech sources by a few degrees at some azimuths (the mean improvement across azimuths was not reported). Dieudonné and Francart (2018) proposed to enhance head shadow ILDs using a fixed beamformer with contralateral attenuation in each ear. The method was tested on normal-hearing listeners simulating bimodal stimulation (i.e., listening with a simulated CI in one ear and a simulated hearing loss in the other ear). RMS localization angle errors improved from 50.5° without the beamformer to 26.8° with the beamformer. While potentially useful for BiCI users, to our knowledge neither method has yet been tested on BiCI users. Therefore, it is not possible to directly compare their benefit with the benefit provided by MOC processing.

Other ILD-enhancement approaches have been specifically designed and tested for BiCI users. For example, Francart et al. (2011) proposed an ILD enhancement algorithm for bimodal CI users that improved individual lateralization scores in a virtual sound field by 4° to 10°. The mean angle error decreased from 28.4° without ILD enhancement to 20.6° with enhancement. Brown (2018) proposed a sound processing strategy intended to provide BiCI with larger than normal ILD cues. Mean angle errors improved from 31.0° without enhancement to 12.8° with ILD enhancement. These ILD-enhancement strategies provide better absolute lateralization scores and larger improvements with respect to the reference condition than the MOC1 strategy (mean angle error = 22.7°, improvement re STD = 2.5°; Fig. 7). We note, however, that the enhancement of ILD cues in the MOC1 strategy is an emergent property of MOC processing rather than an intended effect (Lopez-Poveda, 2015). Furthermore, MOC processing distinguishes itself from the ILD-enhancement methods just described in that it is computationally simpler and requires little streaming between the pair of CIs. Both these characteristics make the MOC strategy suitable for implementation in clinical devices.

4.5. A final remark: which MOC strategy?

In addition to improving sound source lateralization scores, the MOC1 strategy can also improve the intelligibility of speech when the target source is presented in competition with another talker (Lopez-Poveda et al., 2017) or with a source of steady-state noise (Lopez-Poveda et al., 2016a). The MOC1 strategy, however, has potential drawbacks: (1) it can reduce the speech information in the ear opposite to the target source (i.e., the ear with the worse acoustic SNR), which could potentially hinder intelligibility in unilateral listening when the implant ear has the worse acoustic signal-to-noise ratio; and (2) the mutual inhibition between the pair of processors can decrease the overall stimulation levels and thus audibility, which could hinder intelligibility in bilateral or unilateral listening when the two CIs (or processors) have input signals with identical levels. (Note that the latter drawback is less of a concern in realistic listening conditions because any asymmetrical placement of the CI microphones would suffice for the levels of the input signals to be different.) The MOC2 and/or the MOC3 strategy could overcome the two drawbacks (Lopez-Poveda and Eustaquio-Martín, 2018) but their control of compression is too slow to enhance the ILD for brief sounds such as those employed here (200 ms). As discussed above, however, the MOC2 strategy and to some extent also the MOC3 strategy could provide ILDs closer to those of the MOC1 strategy for longer stimuli. Therefore, all three strategies hold potential for improving some aspects of the hearing of BiCI users. Research is ongoing to elucidate which implementation and parameters provide a greater overall benefit for the patient.

5. Conclusions

1. Compared to a STD strategy involving two independently functioning CIS processors with fixed back-end compression, the MOC strategy with fast control of compression and greater inhibition in the higher frequency channels (MOC1), slightly improved the localization of wideband (125–6000 Hz) noise bursts in a virtual horizontal plane.
2. MOC implementations that involved slower control of compression, and/or slightly greater inhibition in the lower than in the higher frequency channels (MOC2 and MOC3 strategies) also provided larger ILDs that the STD strategy for sufficiently long stimuli (>1 s). However, for the shorter (200-ms) noise bursts employed here, the localization performance with these strategies was not significantly different from that with the STD strategy.

3. The localization improvements observed for the MOC1 strategy are probably due to this strategy providing larger and less ambiguous ILDs.

Acknowledgements

We thank Dr. Tim Brochier and two anonymous reviewers for comments on previous versions of this manuscript and Dr. Otto Peter from the University of Innsbruck for equipment and technical support. Work supported by MED-EL GmbH, by the Spanish Ministry of Economy and Competitiveness (grant BFU2015-65376-P to EAL-P), and by the European Regional Development Fund. Author BSW is a consultant for MED-EL GmbH, as its Chief Strategy Advisor.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heares.2019.05.004.

References

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., 2015. Contralateral efferent suppression of human hearing sensitivity. Front. Syst. Neurosci. 8, 251.

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

Bacon, S.P., Takahashi, G.A., 1992. Overshoot in normal-hearing and hearing-impaired subjects. J. Acoust. Soc. Am. 91, 2865–2871.

Blauert, J., 1997. Spatial Hearing: the Psychophysics of Human Sound Localization, second ed. MIT Press, Cambridge, MA.

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

Boyle, P.J., Büchner, A., Stone, M.A., Lenarz, T., Moore, B.C.J., 2009. Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants. Int. J. Audiol. 48, 211–221.

Brown, C., 2018. Corrective binaural processing for bilateral cochlear implant patients. PLoS One 13, e0187965.

Carlyon, R.P., White, L.J., 1992. Effect of signal frequency and masker level on the frequency regions responsible for the overshoot effect. J. Acoust. Soc. Am. 91, 1034–1041.

Ching, T.Y.C., van Wanrooy, E., Dillon, H., 2007. Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review. Trends Amplif. 11, 161–192.

Dorman, M.F., Loiselle, L., Stohl, J., Yost, W.A., Spahr, A., Brown, C., Cook, S., 2014. Interaural level differences and sound source localization for bilateral cochlear implant patients. Ear Hear. 35 (6), 633–640.

Dorman, M.F., Loiselle, L.H., Cook, S.J., Yost, W.A., Gifford, R.H., 2016. Sound source localization by normal hearing listeners, hearing-impaired listeners and cochlear implant listeners. Audiol. Neuro. Otol. 21 (3), 127–131.

Dieudonné, B., Francart, T., 2018. Head shadow enhancement with low-frequency beamforming improves sound localization and speech perception for simulated bimodal listeners. Hear. Res. 363, 78–84.

Francart, T., Lenssen, A., Wouters, J., 2011. Enhancement of interaural level differences improves sound localization in bimodal hearing. J. Acoust. Soc. Am. 130 (5), 2817–2826.

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.

Grantham, D.W., Ashmead, D.H., Ricketts, T.A., Labadie, R.F., Haynes, D.S., 2007. Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. Ear Hear. 28, 524–541.

Jones, H., Kan, A., Litovsky, R.Y., 2014. Comparing sound localization deficits in bilateral cochlear-implant users and vocoder stimulations with normal-hearing listeners. Trends Hear 18.

Kan, A., Litovsky, R.Y., 2015. Binaural hearing with electrical stimulation. Hear. Res. 322, 127–137.

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.

Laback, B., Pok, S.M., Baumgartner, W.D., Deutsch, W.A., Schmid, K., 2004. Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors. Ear Hear. 25, 488–500.

Lilaonitkul, W., Guinan, J.J., 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.

Lopez-Poveda, E.A., 1996. The Physical Origin and Physiological Coding of Pinna-Based Spectral Cues. Doctoral Thesis. Loughborough University, United Kingdom.

Lopez-Poveda, E.A., 2015. Sound Enhancement for Cochlear Implants. Patent Application PCT/EP2015/059230.

Lopez-Poveda EA. 2017. Medial olivocochlear reflex sound coding with bandwidth normalization. US Patent application nº: 62/525902.

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

Lopez-Poveda, E.A., Eustaquio-Martín, A., 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.

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., 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.

Majdak, P., Goupell, M.J., Laback, B., 2011. Two-dimensional localization of virtual sound sources in cochlear-implant listeners. Ear Hear. 32, 198–208.

Majdak, P., Masiero, B., Fels, J., 2013. Sound localization in individualized and non-individualized crosstalk cancellation systems. J. Acoust. Soc. Am. 133, 2055–2068.

Marmel, F., Marrufo-Pérez, M.I., Heeren, J., Ewert, J., Lopez-Poveda, E.A., 2018. Effect of sound level on virtual and free-field localization of brief sounds in the anterior median plane. Hear. Res. 365, 28–35.

Moore, B.C., Kolarik, A., Stone, M.A., Lee, Y.W., 2016. Evaluation of a method for enhancing interaural level differences at low frequencies. J. Acoust. Soc. Am. 140 (4), 2817–2828.

Nopp, P., Schleich, P., D'Haese, P., 2004. Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants. Ear Hear. 25, 205–214.

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. https://doi.org/10.1177/2331216519831492.

Ricketts, T., Grantham, D.W., D'Haese, P., Edwards, J., Barco, A., 2006. Cochlear implant speech processor placement and compression effects on sensitivity and interaural level difference. J. Am. Acad. Audiol. 17, 133–140.

Seeber, B.U., Fastl, H., 2008. Localization cues with bilateral cochlear implants. J. Acoust. Soc. Am. 123, 1030–1042.

Seeber, B.U., Baumann, U., Fastl, H., 2004. Localization ability with bimodal hearing aids and bilateral cochlear implants. J. Acoust. Soc. Am. 116, 1698–1709.

Stöbich, S., Zierhofer, C., Hochmair, E.S., 1999. Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy. Ear Hear. 20, 104–1116.

Tyler, R.S., Noble, W., Dunn, C., Witt, S., 2006. Some benefits and limitations of binaural cochlear implants and our ability to measure them. Int. J. Audiol. 45, S113–S119.

van Hoesel, R.J., 2004. Exploring the benefits of bilateral cochlear implants. Audiol Neurotol 9, 234–246.

van Hoesel, R.J., Tyler, R.S., 2003. Speech perception, localization, and lateralization with bilateral cochlear implants. J. Acoust. Soc. Am. 113 (3), 1617–1630.

Verschuur, C.A., Lutman, M.E., Ramsden, R., Greenham, P., O'Driscoll, M., 2005. Auditory localization abilities in bilateral cochlear implant recipients. Otol. Neurotol. 26, 965–971.

Wenzel, E.M., Arruda, M., Kistler, D.J., Wightman, F.L., 1993. Localization using nonindividualized head-related transfer functions. J. Acoust. Soc. Am. 94 (1), 111–123.

Wiggins, I.M., Seeber, B.U., 2011. Dynamic-range compression affects the lateral position of sounds. J. Acoust. Soc. Am. 130, 3939–3953.

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.

Yost, W.A., Dye, R.H., 1988. Discrimination of interaural differences of level as a function of frequency. J. Acoust. Soc. Am. 83, 1846–1851.

Zeng, F.G., 2004. Compression and cochlear implants. In: Compression: from Cochlea to Cochlear Implants. Springer, New York, NY), pp. 184–220.

Appendix 2.

EXTENDED SUMMARY IN SPANISH

Evaluación multidimensional de una estrategia binaural de procesamiento de sonido para implantes cocleares inspirada en el reflejo olivococlear medial

Resumen extendido en español

Tesis Doctoral

María Milagros Jerónimo Fumero

Director

Enrique A. López-Poveda

Instituto de Neurociencias de Castilla y León

Universidad de Salamanca

Septiembre 2020

RESUMEN

Los implantes cocleares (ICs) pueden proporcionar a las personas sordas una audición eficaz mediante estimulación eléctrica directa del nervio auditivo. A pesar del progreso logrado en el diseño y el rendimiento de los ICs, los usuarios de estos dispositivos todavía tienen dificultades para comprender el habla en ambientes ruidosos o para localizar fuentes sonoras, incluso con ICs modernos y bilaterales.

La estrategia MOC es una estrategia binaural de codificación de sonido para ICs inspirada en el control dinámico de la compresión de la membrana basilar que proporciona el reflejo olivococlear medial (MOCR) contralateral en la audición natural. En contraste con el enfoque clínico estándar (STD), que implica usar dos procesadores de sonido funcionalmente independientes y con compresión acústico-eléctrica fija, la estrategia MOC vincula dinámicamente la cantidad de compresión aplicada en cada oído. Esto puede mejorar el reconocimiento de habla en ruido [Lopez-Poveda et al., 2016, Ear Hear 37(3): e138-148]. Aunque prometedora, la estrategia MOC original presenta algunos inconvenientes y sus parámetros no tienen en cuenta aspectos importantes del MOCR natural. El objetivo principal de esta tesis es evaluar experimentalmente los beneficios proporcionados por implementaciones más realistas de la estrategia MOC sobre la inteligibilidad del habla en ruido, la localización de fuentes sonoras y el esfuerzo de escucha.

La tesis consta de cuatro estudios. El primero de ellos se centró en el reconocimiento de habla en ruido. Se midieron umbrales de recepción de verbal (SRTs) para frases inmersas en ruido estacionario, en condiciones de escucha unilateral y bilateral y para múltiples configuraciones espaciales de las fuentes de habla y ruido. Se compararon los SRTs para estímulos procesados a través de la estrategia STD; la estrategia MOC original, con control rápido de la compresión y mayor inhibición en altas que en bajas frecuencias (MOC1); la estrategia MOC1 con un control más lento de la compresión, y, por lo tanto, más parecido al curso temporal de la inhibición del MOCR (MOC2); y la estrategia MOC2 con mayor inhibición en bajas que en altas frecuencias (MOC3) y, por lo tanto, más parecida al MOCR. Descubrimos que la estrategia más realista (MOC3) corrige las deficiencias de la estrategia MOC1 original y proporciona un mejor reconocimiento de habla en ruido. Además, las estrategias MOC2 y MOC3 proporcionaron una ventaja binaural significativa, algo que no ocurrió con las otras estrategias evaluadas.

El segundo estudio se centró en la lateralización de las fuentes de sonido. Se pidió a los usuarios de IC bilateral que localizaran fuentes de ruido en un plano horizontal virtual para estímulos procesados a través de las estrategias STD, MOC1, MOC2 y MOC3. En comparación con la estrategia STD, la estrategia MOC1 mejoró ligeramente la localización de ráfagas de ruido de banda ancha de 200 ms de duración. Las estrategias MOC2 y MOC3 no mejoraron la localización porque los estímulos eran demasiado cortos para activar y desactivar completamente el control contralateral de la compresión, pero en teoría podrían proporcionar beneficios similares a la estrategia MOC1, para estímulos más largos.

El tercer estudio tuvo como objetivo investigar los beneficios de combinar el procesamiento MOC3 con una estrategia de codificación de sonido (denominada FS4) destinada a preservar la estructura temporal fina del sonido en los cuatro canales de frecuencia más apicales. Los

SRTs para frases procesadas a través de la estrategia MOC3-FS4 y una estrategia estándar FS4 (STD-FS4) se compararon en silencio, en ruido estacionario y en ruido fluctuante, para varios niveles de habla, en escucha bilateral y unilateral, y para múltiples configuraciones espaciales de las fuentes de habla y ruido. En general, los SRTs fueron iguales o mejores con la estrategia MOC3-FS4 que con la estrategia STD-FS4.

El cuarto estudio tuvo como objetivo investigar si el esfuerzo de reconocer el habla en ruido es menor o igual con las estrategias MOC que con las estrategias STD. El porcentaje de palabras recordadas y los tiempos de respuesta verbal en una prueba de reconocimiento de palabras se usaron como indicadores del esfuerzo, y se midieron en silencio y en ruido estacionario a +5 dB de relación señal-ruido (SNR) y en el SRT individual para frases en ruido. Los resultados mostraron que los usuarios de IC bilateral experimentaron aproximadamente el mismo esfuerzo con todas las estrategias de procesamiento de sonido.

En conjunto, los hallazgos demuestran que la estrategia binaural MOC, con parámetros realistas del MOCR natural, puede mejorar la localización de las fuentes de sonido y el reconocimiento del habla en ambientes ruidosos sin aumentar el esfuerzo de escucha. Además, demuestran que es posible combinar el procesamiento MOC con técnicas de codificación de audio para ICs de última generación, lo que hace que la estrategia MOC sea un enfoque prometedor para mejorar aún más el rendimiento auditivo de los usuarios de estos dispositivos.

Palabras clave: implante coclear, eferente olivococlear, compresión del rango dinámico, ruido, inteligibilidad del habla, localización del sonido, esfuerzo auditivo, codificación de audio.

ABREVIATURAS Y ACRÓNIMOS

Algunos de los acrónimos empleados en esta tesis se han tomado directamente del inglés.

AGC: control automático de ganancia (del inglés *Automatic Gain Control*)

IC: implante coclear

CIS: muestreo intercalado continuo (del inglés *Continuous Interleaved Sampling*)

dB: decibelios

FS: escala completa (del inglés *Full Scale*)

HINT: prueba de audición en ruido (del inglés *Hearing-in-Noise Test*)

HRTF: funciones de transferencia asociadas a la cabeza (del inglés *Head-Related Transfer Function*)

iFFM: máscara fluctuante femenina internacional (del inglés *International Female Fluctuating Masker*)

ILD: diferencia interauricular de nivel (del inglés *Inter-aural Level Difference*)

KEMAR: maniquí Knowles Electronics para investigación acústica (del inglés *Knowles Electronics Manikin for Acoustics Research*)

MCL: nivel confortable máximo (del inglés *Maximum Comfortable Level*)

MOC: olivococlear medial (del inglés *Medial Olivo-cochlear*)

MOCR: reflejo olivococlear medial (del inglés *Medial Olivo-cochlear Reflex*)

ms: milisegundo

SNR: relación señal-ruido (del inglés *Signal-to-Noise Ratio*)

SPL: nivel de presión sonora (del inglés *Sound Pressure Level*)

SRT: umbral de recepción verbal (del inglés *Speech Reception Threshold*)

SSN: ruido con espectro del habla (del inglés *Speech-Shaped Noise*)

STD: estándar (del inglés *Standard*)

TFS: estructura temporal fina (del inglés *Temporal Fine Structure*)

1.

INTRODUCCIÓN GENERAL

1.1. MOTIVACIÓN

Los implantes cocleares (ICs) proporcionan una audición eficaz a las personas con sordera o hipoacusia severa-profunda mediante la estimulación eléctrica directa del nervio auditivo. A pesar del progreso logrado en el diseño y el rendimiento de estos dispositivos, los usuarios de ICs aún muestran dificultades para comprender el habla en ambientes ruidosos o para localizar las fuentes de sonido, entre otros aspectos de la audición. Para los usuarios de ICs bilaterales, estas dificultades pueden deberse, en parte, a que los dos ICs funcionan de manera independiente el uno del otro; es decir, a que funcionan como dos dispositivos monoaurales. Es frecuente, por ejemplo, que, en los usuarios de IC bilateral, los dos dispositivos empleen guías de electrodos de diferente longitud (y por tanto electrodos anatómicamente desalineados), diferente número de canales de frecuencia, o incluso diferente tasa de estimulación eléctrica. Estos y otros factores pueden distorsionar o degradar la información acústica binaural imprescindible para reconocer el habla en ambientes ruidosos o localizar los sonidos (Litovsky et al., 2012; Jones et al., 2014; Kan, 2018). Parece razonable suponer, por tanto, que el uso de procesadores binaurales de sonido (es decir, el uso de dos procesadores vinculados, acoplados o coordinados) podría mejorar los beneficios de la implantación coclear bilateral con respecto al uso de dos procesadores de sonidos funcionalmente independientes entre sí.

Lopez-Poveda et al. (2016a, 2016b) demostraron que, comparado con el uso de dos ICs convencionales, el uso de una estrategia binaural de procesamiento de sonido inspirada en el reflejo olivococlear medial contralateral (MOCR) (estrategia denominada 'MOC') puede facilitar la inteligibilidad del habla en presencia de otras fuentes de sonido, tanto para los usuarios de IC unilateral como bilateral. Al contrario del procesamiento estándar (STD) clínico actual, en el que los dos ICs funcionan independientemente el uno del otro, la estrategia MOC combina el funcionamiento de los dos procesadores de sonido para modificar dinámicamente la cantidad de compresión acústico-eléctrica aplicada en cada oído. Esto mejora la información del habla en el oído con la mejor relación señal-ruido (SNR) acústica (Lopez-Poveda y Eustaquio-Martín, 2018).

Sin embargo, la implementación original de la estrategia MOC mostraba dos desventajas: (1) reducía la información del habla en el oído con la peor SNR acústica, lo que podría dificultar la inteligibilidad en condiciones de escucha unilateral cuando el oído implantado tiene la peor SNR acústica; y (2) la inhibición mutua entre los dos procesadores disminuía los niveles generales de estimulación y, por lo tanto, la audibilidad, algo que podría dificultar la inteligibilidad en la escucha bilateral o unilateral cuando los micrófonos de cada oído recogen sonidos idénticos.

Además, la implementación y los parámetros originales de la estrategia MOC no tuvieron en cuenta algunos aspectos del MOCR natural, tales como (1) su curso temporal lento de activación y desactivación (Cooper y Guinan, 2003; Backus y Guinan, 2006), y (2) que causa una mayor inhibición de las respuestas de la membrana basilar en las regiones cocleares apicales que en las basales (Lilaonitkul y Guinan, 2009; Aguilar et al., 2013). Por otro lado, una evaluación técnica de la estrategia MOC, mediante simulaciones de inteligibilidad, predijo que el uso de constantes de tiempo más largas para la activación y desactivación de la inhibición contralateral, combinada con una mayor inhibición en los canales de frecuencia más bajos, podría resolver las limitaciones de la estrategia MOC original, y mejorar la información de la señal en el oído con la peor SNR acústica (López-Poveda y Eustaquio-Martín, 2018). En definitiva, las simulaciones sugirieron que una implementación más realista de las características del MOCR natural podría incrementar los beneficios de la estrategia MOC y generalizarlos a un conjunto más amplio de condiciones de escucha con ICs.

1.2. OBJETIVOS

El objetivo principal de esta tesis es confirmar experimentalmente estas predicciones, comparando la audición de los usuarios de ICs con diversas implementaciones de la estrategia MOC, diseñadas para reflejar de manera más o menos realista las características inhibitorias del MOCR natural. Para lograr este objetivo general, se establecieron los siguientes objetivos específicos:

1. Investigar los posibles beneficios del uso de implementaciones más realistas de la estrategia MOC para el reconocimiento del habla en ruido, en condiciones de escucha unilateral y bilateral. Esto incluye investigar las ventajas binaurales proporcionadas por el procesamiento MOC.
2. Verificar experimentalmente si la localización de fuentes de sonido en un plano horizontal virtual es mejor con la estrategia MOC que con la estrategia STD.
3. Investigar los beneficios de combinar un procesamiento MOC realista (denominado estrategia MOC3) con el procesamiento FS4, en relación con el uso del procesamiento FS4 aislado. El procesamiento FS4 está presente en los ICs más modernos del fabricante MED-EL y está destinado a preservar la TFS de los sonidos en los cuatro canales de frecuencia más apicales.
4. Comparar el esfuerzo de escucha en tareas de reconocimiento del habla en ruido para sonidos procesados con las estrategias MOC y STD.

1.3. HIPÓTESIS

La hipótesis general es que los beneficios del procesamiento MOC (en relación con la estrategia STD) serán mayores con implementaciones más realistas de los efectos naturales del MOCR.

Las hipótesis específicas son:

1. En la estrategia MOC, el uso de constantes de tiempo más largas para la activación y desactivación de la inhibición contralateral, combinado con una mayor inhibición en los canales de frecuencia baja que en los de frecuencia alta, resuelve las limitaciones de la implementación MOC original, e incluso mejora la información de la señal en el oído con la peor SNR acústica. Por lo tanto, implementaciones más realistas de la estrategia MOC producirán un mejor rendimiento en tareas de reconocimiento del habla en ruido, en condiciones de escucha unilaterales y bilaterales para diversas configuraciones espaciales de la máscara y la señal.
2. En comparación con la estrategia STD, las estrategias MOC realzan las diferencias interauriculares de nivel (ILDs), mejorando así la localización de estímulos acústicos en el plano horizontal. Por lo tanto, el rendimiento en una tarea de lateralización de sonidos será mejor con implementaciones más realistas de la estrategia MOC.
3. La estrategia MOC3, en combinación con el procesamiento de sonido FS4, produce mejores SRTs en ruido que el procesamiento FS4 por sí solo.
4. Debido a que las estrategias MOC facilitan el reconocimiento del habla en ruido, escuchar en ruido con estas estrategias requiere el mismo o menor esfuerzo que escuchar con la estrategia STD.

2.

MÉTODOS GENERALES

2.1. PARTICIPANTES

En los estudios participaron 20 usuarios de IC bilateral y dos usuarios bimodales (con IC en un oído y audífono en el otro). Todos los ICs eran de la marca MED-EL. Las pruebas de evaluación se distribuyeron en dos protocolos experimentales. Las **Tablas 3.1** y **3.2** muestran los datos de los participantes del primer y segundo protocolo, respectivamente.

2.2. ESTRATEGIAS DE PROCESAMIENTO DE SONIDO

2.2.1. Estrategia STD

Se evaluaron dos estrategias estándar. Una de ellas, que denominaremos STD, estaba basada en la estrategia de muestreo intercalado continuo (o estrategia CIS, del inglés *Continuous Interleaved Sampling*) (Wilson et al., 1991). Esta estrategia se implementó sin control automático de ganancia (AGC) (**Fig. 2.4**). La otra estrategia estándar, que denominaremos STD-FS4, se basó en la estrategia FS4 de MED-EL, diseñada para preservar la TFS de los sonidos en los cuatro canales de frecuencia más apicales. La estrategia STD-FS4 se implementó con AGC vinculado, es decir, que los AGCs de los dos oídos aplicaron una misma ganancia e igual a la menor de ambos oídos.

En las dos estrategias estándar, los dos procesadores de sonido funcionaban independientemente el uno del otro, y ambos tenían mapas de compresión acústico-eléctrica fijos; es decir, el valor del parámetro *c* en la Ec. (3.1) se mantuvo constante. Estas estrategias fueron similares a las empleadas por los participantes en sus dispositivos clínicos, excepto por el uso del AGC vinculado en la estrategia STD-FS4.

El número de filtros en los bancos de filtros fue idéntico al número mínimo de electrodos activos entre los procesadores izquierdo y derecho (**Tablas 3.1 y 3.2**), e igual para los procesadores del oído izquierdo y derecho.

2.2.2. Estrategias MOC

Los procesadores MOC fueron similares a los procesadores STD, excepto que el valor del parámetro de compresión [*c* en la Ec. (3.1)] en cada canal de frecuencia variaba dinámicamente dependiendo del nivel de salida, ponderado en el tiempo, del canal de frecuencia correspondiente en el procesador contralateral (**Fig. 2.6**). La relación entre el valor de *c* y el nivel de salida contralateral fue tal que cuanto mayor era el nivel de salida, menor era el valor de *c* (**Fig. 2.7**) (Lopez-Poveda et al., 2016a, 2017).

El objetivo principal de esta tesis es evaluar diversas implementaciones de la estrategia MOC, diseñadas para reflejar de manera más o menos realista las características inhibitorias del MOCR natural. En concreto, se evaluaron cuatro implementaciones diferentes de la estrategia MOC:

- MOC1. Es la estrategia MOC implementada y evaluada originalmente (Lopez-Poveda et al., 2016b, 2017), con constantes de tiempo rápidas (τ_a = τ_b = 2 ms), y con mayor inhibición en los canales de frecuencia más alta que en los de frecuencia más baja.
- MOC2. Se trata de una estrategia MOC1 con constantes de tiempo más lentas (τ_a = 2 ms, τ_b = 300 ms) y, por tanto, más parecidas a las del curso temporal de activación y desactivación del MOCR contralateral natural (Backus y Guinan, 2006).
- MOC3. Es una estrategia MOC2 con normalización de ancho de banda (Ec. 3.2), para simular una mayor inhibición en los canales de frecuencia apicales que en los basales, tal y como ocurre con el MOCR contralateral natural (Lilaonitkul y Guinan, 2009b).
- MOC3-FS4. Esta es una estrategia MOC3 con AGC vinculado binauralmente y con procesamiento FS4 para preservar la TFS en los cuatro canales de frecuencia más apicales.

Es importante destacar que las estrategias MOC1, MOC2 y MOC3 tenían la estrategia STD como referencia (protocolo 1), mientras que la estrategia MOC3-FS4 tenía la estrategia STD-FS4 como referencia (protocolo 2).

2.3. AJUSTE Y BALANCE DEL NIVEL DE SONORIDAD

Antes de la realización de las pruebas, se midieron los niveles de corriente eléctrica correspondientes al nivel máximo confortable (MCL, del inglés *Maximum Comfortable Level*), utilizando el método de ajuste. Los niveles mínimos de estimulación (umbrales) se establecieron en valores medidos individualmente, o en 0%, 5% o el 10% de los valores de MCL (Boyd, 2006) (**Tablas 3.1 y 3.2**). Los volúmenes de cada procesador se establecieron utilizando las estrategias STD o STD-FS4 para asegurar que los sonidos se percibían cómodamente e igual de fuertes en ambos oídos, y que una frase filtrada con funciones de transferencia de la cabeza (HRTFs) para 0° de elevación y 0° azimut se percibía en el centro de la cabeza. En los experimentos realizados en el protocolo 1 (dirigidos a comparar las estrategias STD, MOC1, MOC2 y MOC3), los umbrales, los MCL y los volúmenes de cada participante fueron iguales para todos los procesadores y constantes durante todas las pruebas. En los experimentos realizados en el protocolo 2 (dirigidos a comparar las estrategias STD-FS4 y MOC3-FS4), los umbrales y los MCL de cada participante también fueron iguales constantes durante todas las pruebas. Sin embargo, los volúmenes se ajustaron independientemente para cada uno de los dos procesadores con el fin de compensar la posible reducción de sonoridad asociada a la inhibición contralateral presente en la estrategia MOC3-FS4 relativa a la STD-FS4.

2.4. ACÚSTICA VIRTUAL

Para el primer protocolo, la escucha en campo libre se simuló filtrando frases grabadas y almacenadas digitalmente a través de HRTFs ecualizadas de campo difuso, para un maniquí de investigación acústica (KEMAR), y para altavoces situados a 1 m del centro de la cabeza del maniquí (Gardner y Martin, 1995). Para el segundo protocolo, las configuraciones espaciales se consiguieron filtrando las frases a través de HRTFs para el procesador Opus 3 (o3) de MED-EL. Estas últimas fueron proporcionadas por MED-EL. En todos los casos, se simularon situaciones de escucha en las que todas las fuentes de habla y ruido estaban a la altura de los ojos (ángulo de elevación de 0°). Las configuraciones espaciales se eligieron de modo que la fuente de habla siempre estuviera en frente del oyente, o hacia el mejor oído de cada participante (identificado subjetivamente por el propio participante). A menos que se indique lo contrario, las pruebas de audición unilateral se realizaron empleando el mejor oído.

2.5. COVENCIONES ANGULARES

A lo largo de esta tesis, las configuraciones espaciales de los estímulos se expresan como S_XN_Y, donde X e Y indican los ángulos azimutales (en grados) de las fuentes de habla (S) y ruido (N), respectivamente, con 0° indicando una fuente frente al oyente, y valores positivos y negativos indicando ángulos de azimut a la derecha y a la izquierda del plano sagital, respectivamente.

2.6. EQUIPAMIENTO

Se utilizó software a medida (en entorno Matlab R2014a y R2015b, The Mathworks, Inc.), para realizar tanto el procesamiento de señales como los experimentos, incluida la presentación de los estímulos eléctricos. Los estímulos se generaron digitalmente, y se procesaron a través de la estrategia de procesamiento correspondiente. Los patrones de estimulación eléctrica resultantes se transmitieron a los receptores/estimuladores implantados en cada participante utilizando la interfaz de investigación Box 2 (comúnmente conocida como RIB2) del Departamento de Física Iónica y Física Aplicada de la Universidad de Innsbruck (Innsbruck, Austria).

Debido a que todos los estímulos se entregaron directamente al implante del participante, no fue necesario realizar las pruebas en salas aisladas acústicamente. Las pruebas se realizaron en una habitación normal, con el experimentador sentado frente al participante. El experimentador controlaba el software experimental y calificaba las respuestas de los participantes. Todas las pruebas se realizaron en formato "doble ciego", de modo que ni el experimentador ni el participante conocían la estrategia de procesamiento de sonido que se estaba evaluando en cada momento.

2.7. ANÁLISIS ESTADÍSTICOS

Todos los análisis estadísticos se realizaron con IBM SPSS Statistics versión 23.

3.

INTELIGIBILIDAD DEL HABLA CON IMPLEMENTACIONES MÁS REALISTAS DE LA ESTRATEGIA MOC[1]

3.1. INTRODUCCIÓN

Lopez-Poveda et al. (2016b, 2017) demostraron que los usuarios de IC bilateral presentan un mejor reconocimiento del habla en ruido con la estrategia MOC que con la STD. Sin embargo, estos estudios se limitaron a una implementación de la estrategia MOC con constantes de tiempo cortas (2 ms) para la activación y desactivación de la inhibición contralateral. Dicha implementación y parámetros no tuvieron en cuenta algunos aspectos del MOCR natural. Además, utilizando un modelo de inteligibilidad del habla, López-Poveda y Eustaquio-Martín (2018) predijeron que el uso de constantes de tiempo más largas para la activación y desactivación de la inhibición contralateral, combinado con una inhibición comparativamente mayor en los canales de frecuencia baja que en los de frecuencia alta, puede resolver las deficiencias del procesamiento MOC e incluso mejorar la información de la señal en el oído con la peor SNR acústica.

El objetivo principal de este estudio fue confirmar experimentalmente estas predicciones en usuarios de IC. Un segundo objetivo fue investigar la posible ventaja binaural proporcionada por el procesamiento MOC.

3.2. MÉTODOS

En este estudio participaron ocho usuarios de IC bilateral y dos usuarios de IC unilateral. Los SRTs se midieron para frases [versión en castellano (Huarte, 2008) del *hearing-in-noise test* (HINT) (Nilsson et al., 1994)], inmersas en ruido estacionario (ruido con espectro del habla, SSN), en escucha unilateral y bilateral, y para múltiples configuraciones espaciales de las fuentes de habla y ruido. Los SRTs se midieron usando un nivel de habla fijo (–20 dB FS) y variando el nivel de ruido de forma adaptativa hasta conseguir la relación señal-ruido (SNR) a la que el participante reconocía el 50% de las frases que se le presentaban. Como referencia, el nivel de habla de –20 dB FS corresponde aproximadamente a 70 dB SPL en los

---

[1] Este capítulo se basa en el artículo: Lopez-Poveda EA, Eustaquio-Martín A, Fumero MJ, Gorospe JM, Polo R, Gutiérrez Revilla A, Schatzer R, Nopp P, Stohl JS. (**2020**). Speech-in-noise recognition with more realistic implementations of binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. *Ear and Hearing*. https://doi.org/10.1097/aud.0000000000000880

procesadores de audio de IC clínicos de MED-EL. Los SRTs se midieron con la estrategia STD, la estrategia MOC 'original' rápida (MOC1), una estrategia MOC más lenta (MOC2) y una estrategia MOC más lenta con mayor inhibición contralateral en los canales de frecuencia baja que en los de frecuencia alta (MOC3).

3.3. RESULTADOS

En la escucha unilateral y cuando el oído implantado tenía la peor SNR, los SRTs fueron, en promedio, 4 dB peores para la estrategia MOC1 que para la STD (como se esperaba), pero fueron iguales o mejores para las estrategias MOC2 o MOC3 que para la estrategia STD. Por otro lado, en promedio, en escucha bilateral los SRTs fueron 1.6 dB mejores para la estrategia MOC3 que para la STD en todas las configuraciones espaciales evaluadas. En general, todas las estrategias produjeron SRTs significativamente mejores para fuentes de habla y ruido espacialmente separadas, que cuando estaban coubicadas. También observamos una ventaja binaural estadísticamente significativa (es decir, mejores SRTs en escucha bilateral que en unilateral) para las estrategias MOC2 y MOC3, pero no para las estrategias STD y MOC1.

3.4. DISCUSIÓN Y CONCLUSIONES

En la escucha unilateral, el rendimiento fue peor con la estrategia MOC1 que con la STD cuando el oído evaluado tenía la peor SNR acústica. Por el contrario, el rendimiento con las estrategias MOC2 y MOC3 fue comparable al rendimiento con la estrategia STD en esas mismas condiciones. Esto confirma experimentalmente una limitación esperada de la estrategia MOC1 (Lopez-Poveda et al., 2016a, 2016b). Confirma, además, la predicción de que esta limitación puede superarse mediante el uso de un control contralateral de la compresión más lento y parecido al curso temporal del MOCR natural (Lopez-Poveda and Eustaquio-Martín, 2018).

En escucha bilateral, el rendimiento fue mejor con la estrategia MOC1 que con la STD para fuentes de habla y ruido separadas espacialmente, pero no para fuentes coubicadas. Sin embargo, el rendimiento con la estrategia MOC3 fue mejor que con la estrategia STD para todas las configuraciones espaciales evaluadas. En promedio, la estrategia MOC3 mejoró los SRTs en 1.6 dB SNR respecto a la estrategia STD. Esto confirma experimentalmente la segunda predicción de que la otra limitación de la estrategia MOC1 (a saber, que los SRTs serían ligeramente peores con la estrategia MOC1 que con la STD para fuentes de habla y ruido coubicadas), se puede superar mediante el uso de un control más lento de la compresión, combinado con mayor inhibición en los canales de frecuencia más bajos que en los más altos.

Todas las estrategias de procesamiento produjeron desenmascaramiento espacial significativo, es decir, mejor inteligibilidad a medida que se separan espacialmente las fuentes de habla y ruido. Además, las estrategias MOC2 y MOC3 proporcionaron una ventaja binaural estadísticamente significativa (es decir, mejores SRTs en ruido en escucha bilateral que unilateral), algo que no ocurrió con las estrategias STD o MOC1.

4.

LATERALIZACIÓN DE FUENTES DE SONIDO VIRTUALES CON UNA ESTRATEGIA BINAURAL DE PROCESAMIENTO DE SONIDO PARA IMPLANTES COCLEARES INSPIRADA EN EL MOCR[2]

4.1. INTRODUCCIÓN

Muchos usuarios de IC bilateral localizan fuentes de sonido con menos precisión que las personas con audición normal (Dorman et al., 2016). Esto puede deberse, en parte, al uso de ICs que funcionan independientemente el uno del otro y con mapas fijos de compresión acústico-eléctrica, lo cual distorsiona y/o reduce las ILDs. Comparado con el uso de dos procesadores de sonido independientes, la estrategia MOC mediante el uso de compresión acoplada binauralmente, puede mejorar las ILDs en cada canal de frecuencia de procesamiento (véase la Fig. 2 en Lopez-Poveda et al., 2016a) y, por lo tanto, las ILDs en general. En este estudio, investigamos los posibles beneficios de la estrategia MOC para localizar fuentes sonoras en el plano horizontal.

4.2. MÉTODOS

En este estudio participaron 12 usuarios de IC bilateral. La tarea consistió en localizar ráfagas de ruido de 200 ms de duración en un plano horizontal virtual, para estímulos procesados a través de las estrategias STD, MOC1, MOC2 y MOC3. La localización se midió en silencio y a un nivel sonoro de –20 dB FS.

4.3. RESULTADOS

La localización fue ligeramente mejor con la estrategia MOC1 que con la estrategia STD. El ángulo de error promedio fue 22.7° con la estrategia MOC1, frente a 25.3° con la estrategia

---

[2] Este capítulo se basa en el artículo: Lopez-Poveda EA, Eustaquio-Martín A, Fumero MJ, Stohl JS, Schatzer R, Nopp P, Wolford RD, Gorospe JM, Polo R, Gutiérrez Revilla MA, Wilson BS. (**2019**). Lateralization of virtual sound sources with a binaural cochlear-implant sound coding strategy inspired by the medial olivocochlear reflex. *Hearing Research* **379**:103-116.

STD (**Fig. 5.6**). Sin embargo, el ángulo de error de localización no fue estadísticamente diferente con las estrategias STD, MOC2 y MOC3.

4.4. DISCUSIÓN Y CONCLUSIONES

El procesamiento MOC realza las ILDs y produce funciones ILD-versus-azimut más pronunciadas que la estrategia STD, siempre que el estímulo sea más largo que los tiempos de activación y desactivación de la inhibición contralateral. El estímulo empleado en estos experimentos tenía una duración de 200 ms y, por tanto, mayor que el tiempo de activación de la inhibición contralateral de la estrategia MOC1 (2 ms), pero insuficiente para activar totalmente la inhibición contralateral en las estrategias MOC2 y MOC3 (> 300 ms). Por ello, sólo la estrategia MOC1 produjo ILDs mayores, una función ILD-versus-azimut más pronunciada y mejor localización que la estrategia STD. Probablemente, la localización con las estrategias MOC2 y MOC3 también sería mejor que con la STD si los estímulos fueran de mayor duración a los empleados en este estudio.

5.

INTELIGIBILIDAD DEL HABLA CON PROCESADORES MOC Y FS4 COMBINADOS

5.1. INTRODUCCIÓN

Hasta este punto, la estrategia MOC se ha evaluado empleando únicamente estimulación mediante secuencias continuas de pulsos (estrategia CIS convencional), a pesar de que el estándar actual en los procesadores de MED-EL es la estrategia FS4. La estrategia FS4 emplea secuencias de pulsos específicas (diferentes) en cada uno de los cuatro canales de frecuencia más apicales, con el fin de preservar la TFS del estímulo acústico en dichos canales. Además, las evaluaciones de la estrategia MOC realizadas hasta ahora se han restringido a un único nivel de señal de –20 dB FS y no han tenido en cuenta que, en general, los procesadores de audio clínicos incorporan AGC. Para valorar mejor el posible beneficio del procesamiento MOC en una eventual implementación de la estrategia en dispositivos comerciales MED-EL, sería apropiado comparar la inteligibilidad del habla en ruido con y sin procesamiento MOC, en combinación con el procesamiento AGC y FS4, para un mayor rango de niveles de estímulo y tipos de enmascaradores. Ese fue el objetivo del estudio descrito en este capítulo.

5.2. MÉTODOS

En este estudio participaron siete usuarios de IC bilateral. La inteligibilidad en ruido se midió y comparó para las estrategias STD-FS4 y MOC3-FS4. Los SRTs se midieron para frases presentadas en silencio, inmersas en ruido fluctuante (máscara fluctuante femenina internacional, iFFM) y ruido estacionario (SSN), para varios niveles de habla (–48, –38 y –28 dB FS), en escucha bilateral y unilateral y para múltiples configuraciones espaciales de la señal y la máscara. Los corpus de frases utilizados en este estudio fueron la versión española del *Oldenburger Sentence Test* (o test matricial) (Hochmuth et al., 2012), y la versión en castellano del HINT (Huarte, 2008).

5.3. RESULTADOS

En general, todos los participantes fueron capaces de reconocer las frases en silencio independientemente del nivel sonoro al que se le presentaron, incluso a –48 dB FS, que fue el nivel de señal más bajo.

En condiciones de escucha unilateral y bilateral, para los dos tipos de máscaras y todos los niveles de señal evaluados, los SRTs medios fueron iguales o mejores con la estrategia MOC3-FS4 que con la STD-FS4 (**Fig. 6.2** a **Fig. 6.6**). En la condición de escucha bilateral, esta mejora fue de 1 dB en promedio, respecto a la estrategia STD y el beneficio fue mayor a –38 dB FS. Sorprendentemente, los SRTs fueron sistemáticamente peores en escucha bilateral que en unilateral, independientemente de la estrategia. Por otro lado, los SRT medios obtenidos con la estrategia MOC3-FS4 tendieron a ser mejores para el ruido estacionario que para el fluctuante.

5.4. DISCUSIÓN Y CONCLUSIONES

En promedio, los SRTs fueron iguales o mejores con la estrategia MOC3-FS4 que con la estrategia STD-FS4 para todos los niveles de habla y máscaras evaluadas. En general, el beneficio de la estrategia MOC3-FS4 fue mayor para el ruido SSN que para la iFFM porque posiblemente la inhibición contralateral fue menor y similar en los dos oídos para la iFFM que para el SSN.

Por otro lado, los SRTs en ruido siempre fueron mejores en condiciones de escucha unilateral que en bilateral (**Fig. 6.7**). El motivo se desconoce. Un posible factor es el aprendizaje, ya que las pruebas de audición bilateral se realizaron el primer día, mientras que las pruebas de audición unilateral se realizaron tras cuatro días de experimentos bilaterales, cuando ya estaba familiarizado con los experimentos y los sonidos percibidos a través de las diferentes estrategias. Sin embargo, puede haber otros factores, además del aprendizaje. En ocasiones, la estimulación bilateral produce un peor reconocimiento del habla que la estimulación unilateral (véase, por ejemplo, Goupell et al., 2018) debido a probablemente a una inadecuada codificación bilateral de estímulo. Curiosamente, en el estudio reportado en el Capítulo 3 encontramos que para los sonidos procesados con las estrategias MOC3 y STD, que no involucraban AGC o FS4 'vinculados', la estimulación bilateral fue mejor o igual a la estimulación unilateral. Esto sugiere que el peor desempeño con estimulación bilateral también podría estar relacionado con el uso de AGC y/o FS4.

6.

ESFUERZO DE ESCUCHA CON VARIAS IMPLEMENTACIOES DE LA ESTRATEGIA MOC

6.1. INTRODUCCIÓN

En los Capítulos 3 y 5, así como en estudios previos (Lopez-Poveda et al., 2016a, 2017), se ha demostrado que la estrategia MOC puede mejorar la inteligibilidad del habla en ruido para los usuarios de IC. Sin embargo, el efecto del procesamiento MOC en el esfuerzo de escucha aún no se ha investigado. Este capítulo tiene como objetivo investigar si la estrategia MOC afecta al esfuerzo de escucha durante el reconocimiento del habla.

6.2. MÉTODOS

En este estudio participaron 13 usuarios de IC bilateral en total. El capítulo integra dos estudios realizados en diferentes momentos. En el Estudio 1, se utilizó el paradigma de doble tarea para comparar el esfuerzo de escucha experimentado por seis usuarios de IC bilateral con la estrategia STD y con diferentes implementaciones de la estrategia MOC (MOC1, MOC2 y MOC3). La tarea principal consistió en reconocer palabras en silencio y en ruido, y la tarea secundaria en recordar esas palabras. Las palabras se presentaron a un nivel fijo de –20 dB FS. Debido a que la estrategia MOC puede facilitar el reconocimiento del habla en ruido, planteamos la hipótesis de que en ruido y a una SNR dada, escuchar con la estrategia MOC requiere el mismo o menor esfuerzo auditivo que escuchar con la estrategia STD. También planteamos la hipótesis de que, en ruido y para condiciones de igual inteligibilidad (es decir, para el habla en el SRT individual), escuchar con la estrategia MOC requiere el mismo esfuerzo que escuchar con la estrategia STD. En definitiva, ambas hipótesis se resumen en que la mejora en el reconocimiento del habla en ruido con la estrategia MOC respecto a la STD no implica que los participantes experimenten más esfuerzo con la estrategia MOC.

En el Estudio 2 utilizamos el mismo paradigma de doble tarea que en el Estudio 1, y además medimos el tiempo de respuesta verbal para comparar el esfuerzo de escucha con las estrategias STD-FS4 y MOC3-FS4, en siete usuarios de IC bilateral. La metodología utilizada en el paradigma de la doble tarea fue la misma que la utilizada en el Estudio 1. En este caso, las palabras se presentaron a –38 dB FS (que corresponde aproximadamente a –20 dB FS con la estrategia STD). Los tiempos de respuesta verbal se cuantificaron como el tiempo transcurrido desde el final del estímulo (una palabra disilábica) y la respuesta del participante (repetición de la palabra). Los tiempos de respuesta se registraron tanto para las respuestas correctas como las incorrectas, y se midieron tres veces por cada condición y el promedio de las tres medidas se consideró como el tiempo de respuesta final. Presumimos que los tiempos de respuesta serían iguales o menores con la estrategia MOC3-FS4 que con la STD-FS4, lo que indicaría que escuchar con la estrategia MOC3-FS4 requiere el mismo o menor esfuerzo que escuchar con la estrategia STD-FS4.

En los dos estudios, se evaluó el esfuerzo para dos configuraciones espaciales de las fuentes de habla y ruido ($S_{15}N_{-15}$ and $S_{60}N_{-60}$), con ruido tipo SSN, y a tres SNRs diferentes: en silencio, a +5 dB SNR y en el SRT individual. Este último corresponde a la SNR a la cada participante reconocía el 50% de las frases en SSN. Los SRTs individuales se tomaron de los estudios descritos en los Capítulos 4 y 6.

6.3. RESULTADOS

En general, en silencio, los participantes recordaron más palabras con la estrategia MOC3 que con las otras estrategias evaluadas. En ruido y a una SNR igual al SRT individual, los participantes recordaron las mismas palabras (en promedio) con todas las estrategias. Tanto el reconocimiento como el recuerdo de palabras tendieron a ser mejores a +5 dB SNR que en silencio o en el SRT individual, independientemente de la configuración espacial.

Por otro lado, las puntuaciones de recuerdo de palabras y los tiempos de respuesta verbal no fueron significativamente diferentes para las estrategias de procesamiento STD-FS4 y MOC3-FS4. Además, observamos una correlación entre el número de palabras recordadas y el tiempo de respuesta verbal. Esto sugiere que estas dos metodologías posiblemente reflejaban cambios en el esfuerzo de escucha, correspondientes a cambios en las demandas de escucha.

6.4. DISCUSIÓN Y CONCLUSIONES

En general, los resultados de los dos estudios indicaron que la proporción de palabras recordadas no fue significativamente diferente para las estrategias MOC con respecto a su correspondiente estrategia STD. Suponiendo que la proporción de palabras recordadas es una medida de esfuerzo, los datos indican que los usuarios de IC bilateral experimentaron cantidades similares de esfuerzo al escuchar con las estrategias MOC y STD. Este resultado es positivo, ya que los participantes estaban más acostumbrados a escuchar el habla procesada a través de una estrategia similar a la STD que a través de las estrategias MOC.

En los dos estudios, tanto el reconocimiento como el recuerdo de palabras tendieron a ser mejores a +5 dB SNR que en silencio o en el SRT individual, independientemente de la configuración espacial y la estrategia. Esto probablemente se debió a que los SRTs fueron generalmente negativos para la mayoría de los participantes (**Tablas 7.1 y 7.2**), y el reconocimiento de palabras es más difícil a SNR más bajos. Además, los participantes tenían más experiencia en la tarea a +5 dB SNR porque esta condición experimental se administró después de la condición en silencio.

La falta de consenso sobre la metodología más apropiada para evaluar el esfuerzo nos motivó a usar dos metodologías diferentes en el Estudio 2 (doble tarea y tiempo de respuesta verbal). En general, no encontramos diferencias significativas en el esfuerzo entre las estrategias STD-FS4 y MOC3-FS4 con ninguna de las metodologías utilizadas. Aun así, las dos metodologías reflejaron, en cierta medida, los cambios esperados en el esfuerzo con los cambios correspondientes en la demanda de escucha (por ejemplo, en la condición más difícil o SRT). Esto sugiere que los dos métodos podrían reflejar un mismo índice de esfuerzo de escucha.

7.

DISCUSIÓN GENERAL

El principal objetivo de esta tesis fue evaluar experimentalmente diversas implementaciones de la estrategia MOC diseñadas para reflejar de manera más realista las características inhibitorias del MOCR natural. La evaluación consistió en medir el rendimiento de los usuarios de IC en varias pruebas de audición con la estrategia MOC, en comparación con la estrategia STD. Las pruebas de evaluación incluyeron el reconocimiento del habla en ruido para múltiples configuraciones espaciales de las fuentes de señal y máscaras, la localización de una fuente de sonido en el plano horizontal virtual, y el esfuerzo de escucha durante el reconocimiento del habla. En la evaluación, se han incluido estrategias MOC y STD destinadas a descartar, así como a preservar la TFS del habla.

En el Capítulo 3 se ha demostrado que el reconocimiento de frases en ruido es, en general, mejor con una implementación de la estrategia MOC que refleja de manera más realista las características del MOCR natural (la estrategia MOC3). Esta estrategia mantuvo los beneficios de la estrategia MOC1 (propuesta originalmente) sobre la estrategia STD para fuentes de habla y ruido espacialmente separadas, y amplió esos beneficios a otras configuraciones espaciales. Además, la estrategia MOC3 proporcionó una ventaja binaural significativa.

En el Capítulo 4 se ha demostrado que, en comparación con la estrategia STD, la estrategia MOC1 (con un control rápido de la compresión y una mayor inhibición en los canales de frecuencias altas), mejora la localización de ráfagas de ruido de banda ancha (125-6000 Hz) breves (200 ms) en un plano horizontal virtual, ya que realza las ILDs generadas por el efecto de sombra acústica de la cabeza. Las implementaciones de la estrategia MOC con un control más lento de la compresión y/o una inhibición ligeramente mayor en los canales de frecuencias bajas (estrategias MOC2 y MOC3) también podrían proporcionar ILDs más grandes que la estrategia STD, pero para estímulos largos (> 300 ms).

En el Capítulo 5 se ha demostrado que los beneficios del procesamiento MOC3 para el reconocimiento de frases en ruido se mantienen cuando la codificación del sonido incluye el procesamiento FS4, que mantiene parcialmente la TFS del habla. Más específicamente, los SRTs para frases en ruido fueron iguales o ligeramente mejores (~1 dB) con la estrategia MOC3-FS4 que con la STD-FS4, en escucha unilateral y bilateral para diferentes niveles de habla y tipos de máscaras.

Por último, en el Capítulo 7 se ha demostrado que reconocer el habla en silencio y en ruido requiere el mismo esfuerzo con las estrategias MOC que con las estrategias STD.

7.1. COMPARACIÓN CON OTROS ALGORITMOS BINAURALES U OTRAS ESTRATEGIAS DE PROCESAMIENTO BINAURAL

Existen otros enfoques de procesamiento binaural de sonido destinados a que el rendimiento auditivo de los usuarios de IC bilateral sea lo más similar posible al de las personas con audición normal (Wiggins y Seeber, 2013; Baumgärtel et al, 2015a, 2015b; Adiloğlu et al., 2015; Spencer et al., 2018; Potts et al., 2019). Sin embargo, las estrategias de (pre)procesamiento binaural suelen requerir el uso de múltiples micrófonos, algoritmos de detección y mejora del habla, y/o hacer suposiciones sobre las características de la señal objetivo y/o los sonidos de interferencia, o su ubicación espacial (Baumgärtel et al., 2015b). Por el contrario, la implementación de la estrategia MOC en un dispositivo requeriría un micrófono por oído, sin suposiciones a priori sobre la identidad o la ubicación de la señal de interés, sin (pre)procesamiento complejo y, probablemente, menos intercambio de datos entre los oídos. Estas características hacen que la estrategia MOC sea adecuada para la implementación en dispositivos clínicos.

Por otro lado, en la medida en que las ILDs producidas por el efecto sombra de la cabeza se pueden reducir mediante el uso de AGC independientes (no vinculados), y las ILDs naturales se pueden restaurar mediante el uso de AGC vinculado (Wiggins y Seeber, 2013), el procesamiento MOC podría proporcionar mayores beneficios cuando se usa en combinación con un AGC vinculado. Las evaluaciones incluidas en esta tesis se realizaron con la implementación del procesamiento MOC en combinación con una estrategia de codificación de sonido CIS. Sin embargo, hasta donde se sabe, todas las estrategias de codificación de sonido actuales incluyen un mapeo acústico-eléctrico en la(s) última(s) etapas del procesamiento (véase, por ejemplo, la Fig. 2 en Wouters et al., 2015). Teóricamente, el procesamiento MOC podría implementarse con cualquier estrategia de codificación de sonido para IC que aún no utilice mapeos dinámicos. De hecho, el estudio descrito en el Capítulo 5 demuestra que el procesamiento MOC3 se puede combinar con el procesamiento FS4, y que cuando se usa en combinación con AGC vinculado, produce SRTs en ruido iguales o mejores que una estrategia STD-FS4. Sin embargo, se necesita más investigación para comprobar los beneficios de combinar el procesamiento MOC con el AGC vinculado y con otras estrategias de codificación de sonido.

7.2. LIMITACIONES

Debido al número limitado de listas presentes en los corpus utilizados en los estudios de inteligibilidad (Capítulos 4 y 6), es probable que los participantes se aprendieran algunas de las frases utilizadas durante las pruebas. Como resultado, los SRTs obtenidos son probablemente mejores de lo que hubieran sido si el material no se hubiera utilizado repetidamente. Sin embargo, es improbable que la reutilización de las frases haya afectado a las diferencias en los SRTs entre las estrategias (o entre las configuraciones espaciales), ya que las medidas se organizaron en bloques, y en cada bloque se incluyeron todas las estrategias de procesamiento y configuraciones espaciales evaluadas en orden aleatorio. Por lo tanto, incluso si los participantes se aprendieron algunas frases, el efecto de este aprendizaje habría sido similar para todas las estrategias y configuraciones espaciales.

Los beneficios del procesamiento MOC son sorprendentes teniendo en cuenta la falta de experiencia de los participantes con los procesadores MOC. La experiencia con la estrategia MOC se limitó a unos pocos días durante las sesiones experimentales. Por lo tanto, se puede suponer que la mejora proporcionada por la estrategia MOC podría ser mayor si los participantes tuvieran la oportunidad de utilizar a diario esta estrategia de procesamiento.

Por otro lado, en el estudio de localización, la duración del estímulo (200 ms) fue menor que el tiempo requerido para la activación (y desactivación) completa de la inhibición contralateral en las estrategias MOC2 o MOC3 (**Fig. 5.3**). Como resultado, las ILDs para los estímulos evaluados fueron probablemente más pequeñas de lo que hubieran sido para estímulos más largos. Por lo tanto, es posible que el uso de estímulos más largos pueda mejorar el rendimiento de localización también con las implementaciones más realistas de la estrategia MOC.

Dada la falta de consenso sobre la mejor medida del esfuerzo de escucha, utilizamos dos metodologías diferentes (la doble tarea y los tiempos de respuesta verbal) para medir el esfuerzo experimentado por los usuarios de IC con la estrategia MOC en una tarea de reconocimiento del habla. La principal limitación del uso de los tiempos de respuesta como medida del esfuerzo es que son muchos los factores que pueden afectar a la velocidad de procesamiento, como, por ejemplo, la edad (Pichora-Fuller et al., 2016). Una segunda limitación es que el tiempo de respuesta no siempre es sensible al esfuerzo de escucha. Por ejemplo, es posible que dedicar voluntariamente un mayor esfuerzo para mantener el rendimiento de la tarea de como resultado tiempos de respuesta más cortos (Bess y Hornsby, 2014). Por otro lado, el paradigma de la doble tarea puede verse afectado por diferencias individuales en aspectos como el compromiso y la motivación en la tarea (Alhanbali et al., 2019). Además, los estudios que utilizan paradigmas de doble tarea han demostrado que estas medidas de comportamiento sufren de imprecisión y sus resultados son difíciles de comparar con otros estudios (Ohlenforst et al., 2017).

7.3. ¿CUÁL ES LA MEJOR ESTRATEGIA MOC?

En conjunto, los resultados de las diferentes evaluaciones experimentales descritas en esta tesis muestran que el procesamiento MOC mejora la localización de las fuentes de sonido y el reconocimiento del habla en ruido sin aumentar el esfuerzo de escucha. Sin embargo, hemos observado que algunas versiones de la estrategia MOC proporcionan un mayor beneficio en comparación con otras, dependiendo de la tarea, la duración del estímulo y el tipo de máscara utilizada.

Además de mejorar la localización de la fuente sonora, la estrategia MOC1 también puede mejorar la inteligibilidad del habla cuando la señal se presenta en competencia con otro hablante (Lopez-Poveda et al., 2017), o con un ruido de estado estacionario (SSN) (Lopez-Poveda et al., 2016a). En la presente tesis hemos observado que las estrategias MOC2 y/o MOC3 podrían superar los inconvenientes de la estrategia MOC1 (Lopez-Poveda y Eustaquio-Martín, 2018), pero su control de la compresión es demasiado lento para mejorar las ILDs de sonidos breves como los empleados en la prueba de localización descrita en el Capítulo 4. No obstante, las estrategias MOC2 y MOC3 podrían proporcionar ILDs más cercanas a las de la estrategia MOC1 para estímulos más largos. La estrategia MOC3 (más parecida al MOCR natural), resolvió las limitaciones de la estrategia MOC1 original y proporcionó, en general, un mejor reconocimiento del habla en ruido. Además, las estrategias MOC2 y MOC3 proporcionaron una ventaja binaural estadísticamente significativa (Capítulo 3). Esto podría deberse a que las implementaciones más realistas de la estrategia MOC (MOC2 y MOC3) mejoraron ligeramente la información del habla en el oído con la peor SNR acústica y/o probablemente transmitieron información binaural más natural [véase la **Fig. 4.3** y Lopez-Poveda y Eustaquio- Martin (2018)].

En conjunto, parece que la inhibición contralateral más rápida (es decir, como en la estrategia MOC1) podría ser ventajosa para localizar sonidos breves y reconocer el habla en competencia con máscaras fluctuantes, mientras que la inhibición contralateral más lenta (como en las estrategias MOC2 y MOC3) podría ser mejor para el reconocimiento del habla en competencia con un ruido estacionario.

En resumen, todas y cada una de las estrategias MOC evaluadas pueden mejorar algunos aspectos de la audición de los usuarios de IC bilateral. Sin embargo, será importante continuar investigando para esclarecer qué implementación y parámetros proporcionan un beneficio general mayor para los usuarios de IC.

7.4. TRABAJO FUTURO

En general, observamos que, en promedio, los beneficios del procesamiento MOC se mantuvieron para cada uno de los participantes a pesar de que no tuvieron la oportunidad de acostumbrarse por completo a las estrategias MOC antes de las pruebas. Esto sugiere que los beneficios del procesamiento MOC podrían aumentar con la práctica y/o el uso continuo de estas estrategias. Por ello, sería importante proporcionar a los participantes la estrategia MOC en formato hardware portátil, para que puedan usarla en su vida diaria y acostumbrarse a ella. Este hardware también permitiría realizar pruebas similares a las descritas en esta tesis, pero en condiciones de escucha de campo libre real, en lugar de simuladas.

Debido a que la audición en entornos naturales es dinámica (es decir, las personas y los objetos son móviles), y el procesamiento MOC también es dinámico, sería interesante evaluar el rendimiento auditivo con las estrategias MOC utilizando sonidos en movimiento, en escenarios de audición presentes en el mundo real.

Dado que la estrategia MOC mejora la segregación de fuentes de sonido separadas espacialmente (véase la Fig.3 de Lopez-Poveda et al., 2016b), también sería interesante evaluar en qué medida el procesamiento MOC puede facilitar el seguimiento de dos conversaciones simultáneas en entornos realistas.

Debido a que las estrategias MOC2 y MOC3 teóricamente mejoran las ILDs para estímulos suficientemente largos, sería interesante realizar el estudio del Capítulo 4 con estímulos suficientemente largos (>300 ms) como para activar y desactivar completamente la inhibición contralateral. Esto podría ayudar a esclarecer qué implementación y parámetros proporcionan un beneficio general mayor para los usuarios de IC. En otras palabras, sería interesante mejorar los parámetros y las implementaciones de las estrategias MOC evaluadas en esta tesis, para verificar si la inhibición contralateral con constantes de tiempo intermedias podría ser beneficiosa para todos los tipos de máscaras. Se necesita más investigación para confirmar esta última posibilidad.

Por último, también valdría la pena evaluar el esfuerzo de escucha durante una tarea de reconocimiento del habla en ruido con las estrategias MOC utilizando métodos fisiológicos objetivos (en lugar de conductuales), como, por ejemplo, la pupilometría.

CONCLUSIONES

Las principales conclusiones de esta tesis son:

(1) En comparación con una estrategia similar a la STD clínica actual, la estrategia MOC1, mejora ligeramente la localización de estímulos cortos (200 ms) en el plano horizontal virtual. Las estrategias MOC2 y MOC3 también proporcionan beneficios teóricos para estímulos suficientemente largos (> 300 ms).

(2) El reconocimiento del habla en ruido es, en general, mejor con la estrategia MOC3. Esta estrategia mantiene los beneficios de la estrategia MOC1 propuesta originalmente, para fuentes de ruido y habla espacialmente separadas, y amplía esos beneficios a configuraciones espaciales adicionales. Además, la estrategia MOC3 proporciona una ventaja binaural significativa, algo que no ocurre con las otras estrategias utilizadas.

(3) La estrategia MOC3 combinada con el procesamiento FS4 (estrategia MOC3-FS4), produce un reconocimiento del habla en ruido igual o mejor que la estrategia STD-FS4 en condiciones de escucha unilateral y bilateral, para un amplio rango de niveles de habla (−28, −38 y −48 dB FS), para múltiples configuraciones espaciales y para ruidos estacionario y fluctuante.

(4) Durante una tarea de reconocimiento de palabras en ruido, los usuarios de IC bilateral experimentan aproximadamente el mismo esfuerzo de escucha independientemente de que los sonidos se procesen a través de estrategias STD o MOC.

(5) En conjunto, los estudios descritos en esta tesis demuestran que el procesamiento MOC puede mejorar la localización de las fuentes de sonido en silencio y el reconocimiento del habla en ruido sin aumentar el esfuerzo de escucha.