Enhanced detection of envelope-following responses for objective fitting of cochlear-implant users

Julian Schott*†, Robin Gransier†, Marc Moonen*, and Jan Wouters†

**KU Leuven, Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Belgium;*
†*KU Leuven, Department of Neurosciences, ExpORL, Belgium.*

E-mail: ...@kuleuven.be

November 2024

Abstract. *Objective*: Objective cochlear-implant (CI) fitting is the process of using neural brain responses to determine the CI stimulation parameters. Electrically evoked auditory steady-state responses (EASSRs) are potentially useful neural responses for this purpose, but are difficult to detect in electroencephalography (EEG) recordings due to the electrical stimulation artifacts of the CI. A recent approach using system identification (SI) to remove stimulation artifacts and detect EASSRs has shown good detection results but has limitations in clinical practice, such as the need for at least two EEG recordings for optimal results. The objective of this study is to investigate a novel stimulation paradigm that is hypothesized to offer superior SI accuracy, and therefore better response detection, and to be beneficial in clinical practice, as only one EEG recording is required. *Approach*: An amplitude-modulated (AM) CI stimulation pulse train is created where the modulation frequency increases in a step-wise manner within the same stimulation epoch (referred to as SWEEP stimulation). An additional stimulation is created by randomly shuffling modulation frequencies of the SWEEP stimulation within the stimulation epoch (referred to as Shuffled-SWEEP stimulation) to further increase the complexity of the envelope of the stimulation pulse train. AM pulse trains with fixed modulation frequency (referred to as conventional AM stimulation), which are used to elicit EASSRs, are also created for comparison. EEG data is collected from four CI users following stimulation with each of the three distinct paradigms. A supra-threshold stimulation condition is used to investigate whether the SWEEP and Shuffled-SWEEP stimulation can elicit envelope-following responses (EFRs) and how these compare to EASSRs, elicited from conventional AM stimulation. A sub-threshold stimulation condition allows the collection of artifact-only EEG data, which is used to compare the SI accuracy on recordings from the SWEEP and the conventional AM stimulation. *Main results*: In all four CI users, neural responses, following the SWEEP stimulation, Shuffled-SWEEP stimulation, and conventional AM stimulation can be detected in the supra-threshold condition. The validation with artifact-only EEG data shows superiority in SI accuracy when using recordings with SWEEP stimulation compared to recordings with conventional AM stimulation. *Significance*: The presented results hold significant implications for two distinct aspects within the field of objective CI fitting. First, being able to accurately identify the response and artifact within the same EEG recording enables the development of effective, online, objective fitting protocols. Second, detecting EFRs following more complex stimulation paradigms offers a potential advancement in the systematic assessment of the temporal envelope processing in CI users.

1. Introduction

Cochlear implants (CIs) are considered the most successful neuroprostheses that are currently in clinical use. They have revolutionized hearing rehabilitation for individuals with severe to profound sensorineural hearing loss. By bypassing the damaged hair cells in the cochlea and electrically stimulating the auditory nerve, they effectively recreate the sensation of hearing. One challenge in the hearing restoration of CI users is the recurring adjustment of the stimulation parameters. This so-called CI fitting is a time-consuming interaction between the CI user and a trained clinician. The quality therefore highly depends on the communication between both parties, the expertise of the clinician, and many other confounding factors. Fitting results are especially subjective and variable in challenging CI user populations such as infants, and can vary greatly over fitting sessions. Consequently, in the past years, an increasing interest in optimal and personalized fitting using objective measures has emerged in hearing rehabilitation research. Here, electrophysiological responses following CI stimulation are recorded, for instance with electroencephalography (EEG), and utilized to adjust the stimulation parameters.

Commonly used neural responses include the electrically evoked stapedius reflex (Gordon et al. (2004), Caner et al. (2007)), the electrically evoked auditory brainstem response (EABR, Brown et al. (2000), Miller et al. (2008)), electrically evoked compound action potential (ECAP, Hughes et al. (2000), Thai-Van et al. (2004), Miller et al. (2008)), cortical auditory evoked potential (CAEP, Távora-Vieira et al. (2018)), and electrically evoked auditory steady-state response (EASSR, Hofmann & Wouters (2010), Hofmann & Wouters (2012)).

EABRs and ECAPs are transient auditory responses following single- or short bursts of CI stimulation pulses. They are relatively simple to detect and can even be recorded with the CI itself (Dillier et al. (2002), Dijk et al. (2007)), making them therefore popular choices in clinical practice. However, stimulation parameters determined from these responses are suboptimal, as the CI stimulation to elicit these responses and the clinically used CI stimulation differ too much in stimulation rate (McKay & Smale (2017), McKay et al. (2013)). Furthermore, CI fitting based on ECAP recordings has lead to poorer speech comprehension scores than conventional subjective fitting (Seyle & Brown (2002)), or to inconclusive results about the relevance of ECAPs (Smoorenburg et al. (2002)).

The CAEP is a cortical response following CI stimulation pulse bursts, that represent speech tokens such as /m/, /g/, /t/, or speech syllables such as /ba/, /da/, /ga/. The presence of CAEPs in EEG recordings may indicate that speech (speech sounds) is perceived and audible to the CI user (Carter et al. (2013)). CI stimulation levels determined from CAEP recordings have shown a good correlation with behaviorally determined stimulation levels (Visram et al. (2015), Mao et al. (2019)) and adjustments to stimulation parameters based on the presence of CAEPs can lead to improvements in speech comprehension scores (Távora-Vieira et al. (2022)). However, CAEPs are often elicited with short stimuli, that cannot fully represent the temporal aspects of daily-life communication.

More appropriate neural responses as objective measure are EASSRs, as the main signal feature of an acoustic input signal that is recorded and transmitted via CI stimulation, is the envelope, i.e., the slow amplitude fluctuations of the acoustic signal (Wouters et al. (2015)). EASSRs belong to the envelope-following responses (EFRs), where the brain activity follows the envelope of the stimulation pulse train. CI stimulation patterns that recreate the amplitude fluctuations of an acoustic input signal lead to good speech understanding performance of CI users in quiet listening conditions (Shannon et al. (1995))(Drullman et al. (1994)). In the case of EASSRs, the stimulation envelope is a sinusoidal signal with fixed modulation frequency and depth, i.e., an amplitude-modulated (AM) stimulation pulse train is used (further referred to as conventional AM stimulation). Recent research using EASSRs has shown that they can be used to assess the temporal envelope encoding in CI users (Luke et al. (2015), Gransier, Carlyon & Wouters (2020)) and that the encoding ability correlates with speech-in-noise understanding performance (Gransier, Luke, Wieringen & Wouters (2020)). Furthermore, EASSR amplitude growth shows a correlation with perceived loudness growth when increasing stimulation levels as seen in Eeckhoutte et al. (2016).

It is therefore hypothesized that using EASSRs to objectively determine CI stimulation parameters is favorable, as the stimulation levels, that elicit EASSRs, ensure information transmission of the temporal modulations of the stimulus, which is essential for speech understanding. This is especially important in the early

development stages of pediatric CI users, where fitting is challenging, but modulation perception is crucial for speech development (Petley et al. (2024)).

The main obstacle in using EASSRs is the presence of stimulation artifacts of the CI, which contaminate the EEG recording. Like the EASSR, the artifact signal follows the envelope of the stimulation and is therefore impossible to distinguish from the neural response without sophisticated artifact removal methods. However, this is difficult due to the continuous, periodic nature of the stimulation and neural response, which intrinsically leads to a timely overlap of both signals. Conventional filtering is also not possible as the neural response and artifact also show overlapping frequency content, i.e., a component at the modulation frequency of the conventional AM stimulation (Deprez, Gransier, Hofmann, van Wieringen, Wouters & Moonen (2017)). Many artifact removal methods for EASSR recordings have been introduced. They include model based approaches (Luke & Wouters (2017)), blind source separation with independent component analysis (Deprez et al. (2018), Alemi et al. (2021)), denoising source separation (Undurraga et al. (2021)), template subtraction (Deprez, Gransier, Hofmann, Wieringen, Wouters & Moonen (2017)), selective downsampling of the EEG recording (Hu & Ewert (2021)) and linear interpolation between artifact-free EEG recording samples (LI, Hofmann & Wouters (2010)). However, only the recently introduced system identification (SI) approach has shown good results for artifact removal and EASSR detection with clinical stimulation settings of the CI, i.e., high pulse rates at 900 pulses per second (pps) and monopolar stimulation, and with every EEG recording channel position (ipsi- and contralateral to the CI) (Schott et al. (2022), Schott et al. (2023)). Good results have been achieved for a large EASSR dataset from different CI users (N = 16) at four different stimulation levels (Schott et al. (2023)). These findings demonstrate the capability of SI to enable the utilization of clinically relevant EASSRs as electrophysiological responses.
The main shortcoming of the SI-based artifact removal as introduced by Schott et al. (2022), is the need for at least two EASSR recordings, each at a different modulation frequency, to create a model that enables a group delay identification, which serves as validation of the presence of a neural response. Using only one EASSR recording with one modulation frequency is also possible (as done in Verwaerde et al. (2024)), but less accurate in neural response and artifact separation, and does not allow a group delay identification. When providing more input-output information to the SI, i.e., more recordings at different modulation frequencies, a better identification of the systems is achieved (Ljung (1999), Nelles (2001)), and it is therefore hypothesized that this leads to a more reliable separation of stimulation artifact and neural response. However, currently, EASSR EEG recordings of approximately 5 min duration (to guarantee sufficiently low background noise levels) are acquired at each individual modulation frequency (Schott et al. (2023), Gransier, Carlyon & Wouters (2020)), which leads to a high accumulated total recording time, which is undesired in clinical practice. Finally, the current way of recording and analyzing data does not allow the implementation of an online detection algorithm, where the presence of a neural response can continuously be checked while EEG is further recorded.

To overcome the described issues, a novel stimulation paradigm for CI users is introduced, which is based on well-established SI methods from other research or engineering fields (Farina (2000), Müller & Massarani (2001)) and which has also already been used in normal-hearing subjects to determine the modulation frequency transfer function (MFTF) (Purcell et al. (2004), Purcell & John (2010), Picton et al. (2003)). This new paradigm uses a sweep over modulation frequencies and can therefore be used to determine the group delay within a single EEG recording, which is assumed to improve the SI accuracy, time efficiency, and allows online detection of the neural response.
In this paper, the potential benefits of incorporating more input-output information, i.e., more modulation frequencies in the stimulation (further referred to as SWEEP stimulation), are analyzed for SI accuracy and neural response detection. This is achieved by creating hybrid EEG data that combines real artifact-only recordings with simulated neural response signals. Further, it is investigated whether EFRs can be elicited in CI users with the SWEEP stimulation and whether they relate to EASSRs. Additionally, the complexity of the SWEEP stimulation is further increased by randomly shuffling modulation frequencies during the stimulation epoch, and it is investigated whether EFRs can also be elicited and detected with this stimulation envelope. This can provide insight into whether the TEMPEST framework, as introduced by Gransier & Wouters (2021), can be applied to CI users as well.

2. Materials and Methods

2.1. SWEEP stimulation

The idea of the SWEEP stimulation is based on general SI procedures, where sinusoidal sweeps are used to efficiently identify the transfer function of unknown systems (Ljung (1999), Müller & Massarani (2001), Nelles (2001)). Applications of sinusoidal sweep recordings

can be found in different fields such as biomedical engineering (Bernardi et al. (2017)), acoustics (Farina (2000)), and others. In audiology, sinusoidal sweeps have been used, among other things, to evaluate temporal auditory acuity (Purcell et al. (2004), Purcell & John (2010)). In Purcell et al. (2004), an acoustic stimulus is generated by modulating a noise or pure-tone carrier signal with a sinusoid, where the modulation frequency first increases linearly and then decreases linearly. The results show that by using this stimulus, the modulation frequency transfer function can efficiently be evaluated and that the response at distinct frequencies is equal to ASSRs at the same frequency.

As commonly known, the temporal acuity of CI users may be impaired. Therefore, it is hypothesized, that a sweep with step-wise modulation frequency increase, i.e., where the modulation frequency is kept constant over a defined time segment after each increase, facilitates phase locking of the auditory nerve fibers to the modulation of the stimulus. The modulation frequency is kept constant over a time segment T_{seg} that is chosen to be $1.024s$, as this stimulus duration is known to elicit EASSRs with conventional AM stimuli, even if the continuous stimulus is interrupted by a silence period (David et al. (2023)) or a change in modulation depth (Gransier, Carlyon & Wouters (2020)). The modulation frequencies are rounded so that only full cycles of the modulation are contained within the time segments of T_{seg}. This way, the instantaneous phase $\phi(t)$ of the modulating signal is continuous, also when the modulation frequency is increased. This is done to avoid jumps in the stimulation levels when increasing the modulation frequency or repeating the sweep, which would otherwise lead to undesirable 'clicks' in the stimulation.

To detect the neural response and assess its statistical significance, i.e., test whether the determined neural response differs from the neural background noise, a one-sample Hotelling T^2 (HT2) test is applied to the variance of the neural response across multiple recordings of repetitions of the stimulation (epochs). The number of used epochs, therefore, determines the duration of one SWEEP stimulation epoch when aiming for a total recording duration of 5.12 minutes, in line with EASSR recording times (Schott et al. (2023), Gransier, Carlyon & Wouters (2020)). This also means, that the number of modulation frequency steps is limited, when $T_{seg} = 1.024s$. For convenience, a sample size of 60 epochs is chosen, which results in a SWEEP stimulation duration of $5.12s$, which means 5 modulation frequencies can be used within one SWEEP, each corresponding to a time segment of $T_{seg} = 1.024s$. In the frequency range of interest from $36Hz$ to $43Hz$, the following modulation frequencies are used:
$f_m^1 = 36.1328125Hz$, $f_m^2 = 38.0859375Hz$,
$f_m^3 = 40.0390625Hz$, $f_m^4 = 41.015625Hz$
and $f_m^5 = 42.96875Hz$.
The resulting instantaneous frequency $f_m(t)$ of the modulating signal $m(t)$ is shown in Fig. 1.

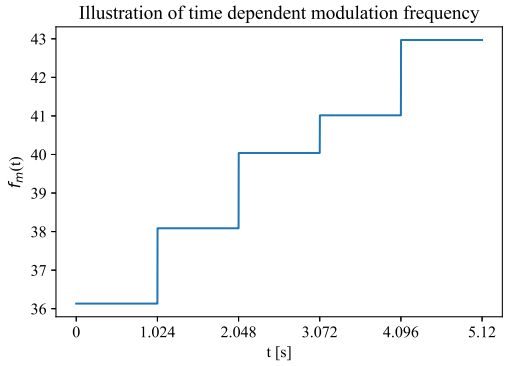


Figure 1. Illustration of time dependent modulation frequency. The modulation frequency $f_m(t)$ increases in discrete steps over the duration of one SWEEP stimulation epoch.

The modulating signal, i.e., the envelope of the CI stimulation pulse train, is then defined as

$$m(t) = \sin(2\pi f_m(t) \cdot t + \frac{\pi}{2}). \quad (1)$$

2.2. Shuffled-SWEEP stimulation

In Gransier & Wouters (2021), the *Temporal Envelope Speech Tracking* (TEMPEST) framework is introduced. This framework can be used to create stimuli to assess the temporal envelope processing in the auditory pathway in a systematic and standardized way. The complexity of the envelope of the stimulus can be varied, from a simple single-frequency envelope where the modulation is kept constant in frequency and depth (e.g. as in equ. (7)), to a speech-like envelope, where modulation depth and frequency continuously vary. The SWEEP stimulation of section 2.1 can be seen as a stimulus from Gransier & Wouters (2021), where the modulation depth stays constant, but the frequencies change over time.

To investigate the effect of a more complex stimulation paradigm than conventional-AM- or the SWEEP stimulation, a TEMPEST-like stimulation for CI users, further referred to as Shuffled-SWEEP, is created. Here, the SWEEP stimulation of section 2.1 is used and the sequence of the instantaneous modulation frequency $f_m(t)$ is randomized. However, it is ensured, that for every modulation frequency in the random

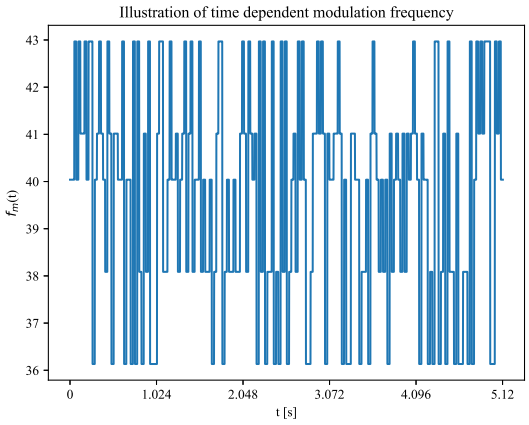


Figure 2. The change of the modulation frequency f_m over the duration of one Shuffled-SWEEP stimulation epoch.

sequence, a full period of duration $T_m = \frac{1}{f_m}$ is created.

$$mod(t) = \sin(2\pi f_m \cdot t + \frac{\pi}{2}) \text{ for t} = 0, \ldots\ T_m \qquad (2)$$

The complete modulating signal m_{SSW}, that is used for the stimulation pulse train, is defined as the concatenation of all full periods

$$m_{SSW} = (mod^1, mod^2, \ldots mod^I), \qquad (3)$$

with $I = T_{seg} \cdot f_m^1 + T_{seg} \cdot f_m^2 + T_{seg} \cdot f_m^3 + T_{seg} \cdot f_m^4 + T_{seg} \cdot f_m^5$, i.e., the total number of periods in the epoch.

By concatenating full periods, a continuous modulation m_{SSW} is again guaranteed. The resulting m_{SSW} is closer related to a speech envelope, where the modulation 'randomly' changes over time, without having confounding effects such as speech understanding influencing the strength of the elicited neural response.

The SWEEP and Shuffled-SWEEP provide equal input information for the SI, as the same modulation frequencies and equal number of full periods per frequency are used, only the sequence of the modulation frequencies randomly changes during one epoch of the Shuffled-SWEEP stimulation as seen in Fig. 2.

2.3. Artifact removal and neural response detection

As discussed in the introduction, the main obstacle in detecting EASSRs (or any other EFR) in EEG recordings, is the presence of CI stimulation artifacts. The SI approach as recently introduced (Schott et al. (2022)) and further improved (Schott et al. (2023)) shows very promising results in the efficacy of separating stimulation artifact and neural response. The method can also be applied to EEG recordings following SWEEP or Shuffled-SWEEP stimulation. The elicited neural responses will hereafter be referred to as EFRs or Shuffled-EFRs, respectively.

For completeness, we repeat the model formulation as defined in Schott et al. (2023).
For the EASSR recordings (following conventional AM stimulation), the recording model remains

$$\underbrace{\begin{bmatrix}\mathbf{z}^{37}\\ \mathbf{z}^{40}\\ \mathbf{z}^{43}\end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix}\mathbf{M}^{37} & \mathbf{U}^{37}\\ \mathbf{M}^{40} & \mathbf{U}^{40}\\ \mathbf{M}^{43} & \mathbf{U}^{43}\end{bmatrix}}_{\mathbf{H}} \cdot \underbrace{\begin{bmatrix}\mathbf{b_{resp}}\\ \mathbf{b_{arti}}\end{bmatrix}}_{\mathbf{x}} + \mathbf{v}, \qquad (4)$$

with recording $\mathbf{z}^{f_m} = \begin{bmatrix} z^{f_m}[0] & z^{f_m}[1] & \ldots & z^{f_m}[K]\end{bmatrix}^\intercal$ being the mean epoch of the pre-processed EEG recording with modulation frequency f_m of the conventional AM stimulation, with a total number of $K+1$ samples. $\mathbf{v}$ is the similarly defined additive noise. The input matrix $\mathbf{M}^{f_m}$ is defined for each EEG recording with modulation frequency f_m as

$$\mathbf{M}^{f_m} = \begin{bmatrix}\mathbf{m}^{f_m}[0]\\ \mathbf{m}^{f_m}[1]\\ \vdots\\ \mathbf{m}^{f_m}[K]\end{bmatrix}, \qquad (5)$$

with

$$\mathbf{m}^{f_m}[k] = \begin{bmatrix} m[k] & m[k-1q] & \ldots & m[k-N_1 q]\end{bmatrix}, \qquad (6)$$

and $m[k]$ being the modulation of the conventional AM stimulation at discrete time index k, with sampling interval T_s of the EEG recording system.

$$m[k] = \sin(2\pi f_m k T_s + \frac{\pi}{2}). \qquad (7)$$

The time-shift parameter q is set to $q = 4194$, as described in Schott et al. (2023). N_1 is the order of the finite-impulse response (FIR) system, which models the EASSR in the EEG recording, and is set to $N_1 = 5$. $\mathbf{U}^{f_m}$ is similarly defined as $\mathbf{M}^{f_m}$ with

$$\mathbf{u}^{f_m}[k] = \begin{bmatrix} u[k] & u[k-1] & \ldots & u[k-N_2]\end{bmatrix}, \qquad (8)$$

using

$$u[k] = A_c \cdot (1 + D \cdot m[k]) \cdot s[k]. \qquad (9)$$

N_2 is the order of the FIR system, which models the stimulation artifact in the EEG recording, and is set to $N_2 = 364$ (Schott et al. (2023)).
$s[k]$ defines the stimulation pulse train with pulse rate f_c, which contains 1's at the sample positions k, which are closest to the stimulation pulse positions and otherwise 0.

$$s[k] = \begin{cases} 1, & \text{for } mod(kT_s, \frac{1}{f_c}) < T_s \\ 0, & \text{otherwise}\end{cases}, \qquad (10)$$

A_c and D are defined by the used stimulation level CL.

$$A_c = \frac{CL + (CL - 50)}{2} \tag{11}$$

and

$$D = \frac{C - (C - 50)}{C + (C - 50)}. \tag{12}$$

For the EFR model, the modulation $m[k]$ is adapted to use the time varying modulation frequency

$$m[k] = \sin(2\pi f_m(kT_s)kT_s + \frac{\pi}{2}), \tag{13}$$

as defined by the SWEEP-stimulation. For the Shuffled-EFR model, $m[k]$ uses m_{SSW} as defined in (3), sampled at discrete time intervals of kT_s.

$$m[k] = m_{SSW}(kT_s). \tag{14}$$

As the SWEEP or Shuffled-SWEEP stimulation use 5 distinct modulation frequencies in $m[k]$, $N1$ is set to $N1 = 9$ for the EFR and Shuffled-EFR models, and $q = 2330$.

For EEG recordings following SWEEP or Shuffled-SWEEP stimulation, the recording model $\mathbf{z}$ reduces to an unstacked model, as no combination of different recordings is required anymore.

$$\mathbf{z} = \underbrace{\begin{bmatrix}\mathbf{M} & \mathbf{U}\end{bmatrix}}_{\mathbf{H}} \cdot \underbrace{\begin{bmatrix}\mathbf{b_{resp}} \\ \mathbf{b_{arti}}\end{bmatrix}}_{\mathbf{x}} + \mathbf{v}. \tag{15}$$

The unknown model coefficients are estimated using the same Tikhnonov regularization as in Schott et al. (2023), and the regularization constant α is equally determined as in Schott et al. (2023) and set to the same value $\alpha = 5000$.

$$\hat{\mathbf{x}} = \begin{bmatrix}\hat{\mathbf{b}}_{\mathbf{resp}} \\ \hat{\mathbf{b}}_{\mathbf{arti}}\end{bmatrix} = (\mathbf{H}^\intercal\mathbf{H} + \alpha\mathbf{\Gamma}^\intercal\mathbf{\Gamma})^{-1}\mathbf{H}^\intercal\mathbf{z}. \tag{16}$$

The response FIR system, used to determine the presence of a neural response in the EEG recordings, is defined as

$$FIR1(z) = \begin{bmatrix}1 & z^{-q} & z^{-2q} & ... & z^{-N_1 q}\end{bmatrix} \cdot \hat{\mathbf{b}}_{\mathbf{resp}}. \tag{17}$$

The group delay of $FIR1$ represents the latency of the neural response and its amplitude is determined as

$$\begin{aligned} A(f_m) &= |FIR1(z = e^{j\omega_m})| \\ &= |\begin{bmatrix}1 & e^{-jq\omega_m} & e^{-j2q\omega_m} & ... & e^{-jN_1 q\omega_m}\end{bmatrix} \cdot \hat{\mathbf{b}}_{\mathbf{resp}}|, \end{aligned} \tag{18}$$

with normalized angular modulation frequency $\omega_m = 2\pi f_m/F_s$.

2.4. Dataset

EEG recordings were collected from four adult, unilateral CI users, all with Cochlear Ltd. devices (CIC4 implants), and at least three years of CI experience. This study was approved by the Medical Ethical Committee of the UZ Leuven Hospital (study number: S65454). All participants gave written informed consent before participation.

The conventional AM stimulation to elicit EASSRs consisted of AM pulse trains with symmetric biphasic cathodic-first pulses, with equal phase width of $25\mu s$ or $37\mu s$, depending on the user's clinical settings, and an interphase gap of $8\mu s$. They were generated using custom software and directly delivered to the implant using the NIC4 research interface platform (Cochlear Ltd.). A clinically relevant pulse rate of $900pps$ was used, and the stimuli were presented between the intracochlear electrode 15 and the extracochlear plate electrode (MP1 mode). Three distinct recordings with respective modulation frequencies of $37Hz$, $40Hz$, and $43Hz$ were used and repeated at two different stimulation levels. A fixed modulation depth of 50 current levels (CL) between the defined stimulation level and the level minus 50 CL was used for the AM pulse trains. The two stimulation levels were defined as the user's C and user's $T - 10CL$ ('Tm10') level. For each user, the T- and C-level were behaviorally determined before the EEG recordings. Here, an AM stimulus with a modulation frequency of $40Hz$, a modulation depth of 50CLs, and a duration of 1s was used. T-and C-levels were determined based on a 7-point categorical loudness scale (i.e. "inaudible", "very soft", "comfortably loud", "loud", "very loud", and "unbearable"). First, the C-level was assessed by gradually increasing CLs from "inaudible" to the most comfortable loudness level which is defined as the current level prior to the transition between "comfortably loud" and "loud" on the categorical loudness scale. Second, the current was gradually decreased to determine the T-level of the modulated pulse train, which was defined as the last current level at which the user could perceive the stimulus before it became "inaudible".

SWEEP and Shuffled-SWEEP stimulation were generated using the same settings as for the conventional AM stimulation. However, instead of creating three distinct stimuli, only one stimulus with a time-varying modulation frequency, as described in 2.1 and 2.2, was created for each stimulation level in each paradigm.

A Biosemi ActiveTwo Hyper-Rate EEG recording system with a sampling rate of 262,144kHz was used to perform the 8-channel recordings. The high sampling

rate of the EEG system allows using other artifact removal methods such as LI, mainly because the cut-off frequency of the front-end low-pass filter of the EEG system is much higher than for common systems. This reduces the temporal 'smearing' of the stimulation pulses in the recording. Consequently, the time duration of the stimulation artifact is shorter than an inter-pulse interval in recordings from contralateral EEG channels, which enables artifact removal with LI in these channels as seen in Gransier, Carlyon & Wouters (2020) and Schott et al. (2023). The high sampling rate therefore provides more signal processing options of the recordings, but it is not a prerequisite for the success of the SI method.
Six recording electrodes were placed according to the international 10-20 system at $P9$, $P10$, Iz, Cz, Fz, Fpz, and two extra electrodes at the left and right mastoids respectively (MaL, MaR). In cases where the EEG recording electrode was located at the CI transmitter/ receiver coil position, the nearest neighbor electrode position was used instead.

A total signal of 5.12min duration was recorded per stimulation paradigm, resulting in 300 epochs per recording of conventional AM stimulation, and 60 epochs for recordings of SWEEP and Shuffled-SWEEP stimulation. The EEG signals were low-pass filtered with the system's built-in analog third-order antialiasing filter having a -3 dB point at 50 kHz. All off-line processing was performed using Matlab R2020a. Off-line referencing with Fz was performed and detrending with a second-order polynomial was applied, using Matlab's built-in *detrend* function. The signals were then, based on the stimulation triggers, split into epochs of their respective duration, and 5% of the epochs with the highest peak-to-peak amplitudes were rejected. The mean epoch for MaR, MaL, $P9$, $P10$, and Iz was used for the further analysis.

2.5. Validation with hybrid EEG data

To investigate the hypothesized benefit of using more modulation frequencies in the SWEEP stimulation on the SI accuracy and hence response detection, hybrid EEG data is created as a combination of real artifact-only EEG recordings and simulated responses. By using this hybrid data, various response systems, i.e., different response latencies and amplitudes, can be simulated, allowing for validation of the detection since the underlying ground truth is simulated and therefore known. Using artifact-only EEG recordings allows for the creation of realistic testing conditions.
The hybrid data is created by using EEG recordings, where the CLs are below the user's T-level ('Tm10' condition). This way, no response is elicited, and only stimulation artifact is present at the modulation frequencies in the EEG recording. The simulated responses are created by using the envelope of the stimulation, i.e., (7) for the conventional AM and (13) for SWEEP stimulation respectively, and digitally filtering it with a simulated FIR system defined as

$$H(z) = \sigma \cdot \ z^{-D} \tag{19}$$

with a simulated group delay in $D = \lfloor delay \cdot F_s \rceil$ samples at sampling frequency F_s, and simulated scaling σ. For instance an EASSR with $40ms$ latency and amplitude of $0.3\mu V$ can be simulated by setting $D = \lfloor 0.04s \cdot 262144Hz \rceil = 10485$ and $\sigma = 0.3$.
The simulated group delays, further referred to as response latencies, range from 0 to 100ms in steps of 1ms, and per simulated response latency, different scalings σ are used. The σ-s are determined from a range of simulated response SNRs, with a fixed hypothetical background noise level v of $v = 50nV$.

$$\sigma = 10^{\frac{SNR}{20}} \cdot 50 \cdot 10^{-9} \tag{20}$$

The simulated SNRs range from -10dB to 20dB in steps of 1dB.
All together, responses from 31x101 simulated response systems are created and added to artifact-only EEG recordings from ipsilateral EEG channels, to ensure sufficient artifact contamination and hence challenging testing conditions.
The conventional AM stimulation hybrid data is further differentiated and validated by either using only one recording of one modulation frequency ($40Hz$), or all three recordings ($37Hz$, $40Hz$ and $43Hz$). For a fair comparison of the simulated conditions, the number of epochs of the recordings of the conventional AM stimulation are reduced, so that equal total recording lengths between conditions are compared. For instance, for the stacked recording model of 3 f_ms, only 1/3 of the epochs per 5.12min recording are used to calculate the mean epoch, which is being analyzed. This way, the stacked recording uses the same amount of data ($3 \cdot \frac{5.12min}{3}$) as the SWEEP recording. The included epochs are randomly selected.

Hybrid data is created for only three of the four subjects, as for one subject technical issues were experienced during the acquisition of the artifact-only EEG recording.
SI is used to perform artifact removal and neural response detection, and the accuracy is validated by checking the true positive rate (TPR) and F1 score. The TPR is the number of detected responses (true positives) over detectable responses, and is used as a measure of sensitivity. The F1 score is the harmonic mean of TPR and precision, of which precision is the ratio of true positives and predicted positives (true and false). F1 is a measure of predictive performance.
Furthermore, the neural response detection is also

validated by analyzing the accuracy of the estimated response latencies and amplitudes.

3. Results

3.1. Hybrid EEG data results

The TPR and F1 score as functions of the simulated response latencies are shown in Fig. 3 and Fig. 4 respectively. For each simulated response latency, SI is applied on the hybrid data of all simulated scalings σ, and HT2 is used to determine whether the identified response is significant, i.e., detected. Simulated responses with SNRs of 6dB or higher, with recording background noise level v, should be, after successful artifact removal, detectable above the noise floor and therefore be statistically significant when applying the HT2 test.

Each plot-line represents the results for a particular stimulation paradigm, i.e., blue for one recording with conventional AM stimulation, green for the combination of three recordings of conventional AM stimulation, and red for the recording of SWEEP stimulation. The following observations can be made from the figures 3 and 4.

One, when using only one recording of conventional AM stimulation, the TPR repeatedly drops to values close to 0%, meaning that no response is detected, even when the simulated SNRs are above 6dB. When using three recordings of conventional AM stimulation or one recording of SWEEP stimulation, the steep drops in TPR disappear, however, fluctuations in TPR remain over different simulated response latencies.

Two, the more modulation frequencies are used in the analyzed recording model, the better the balance between detection accuracy and detection precision, which can be seen from the F1 scores. For one f_m, the mean value is $\bar{F}1 = 0.75$, for three f_m $\bar{F}1 = 0.85$, and for the sweep $\bar{F}1 = 0.93$.

Furthermore, the SI accuracy can also be analyzed by checking the determined response latency. Here, only simulation conditions with response SNRs of 6dB or higher are evaluated, as only these responses should be detectable/ identifiable. The detected response latencies for these are averaged and plotted for each stimulation paradigm as a function of the simulated response latency in Fig. 5. The following observations can be made from the figure.

One, when using only one recording of conventional AM stimulation, no group delay can be determined with the response model, and hence the response latency cannot be detected.

Two, when using three recordings of conventional AM stimulation or one recording of SWEEP stimulation, the detected latency follows the simulated latency. The accuracy for hybrid data with SWEEP stimulation is the best, as seen from the smallest variation and smallest deviation from the simulated response latencies. Three, lower simulated latencies are less accurately detected.

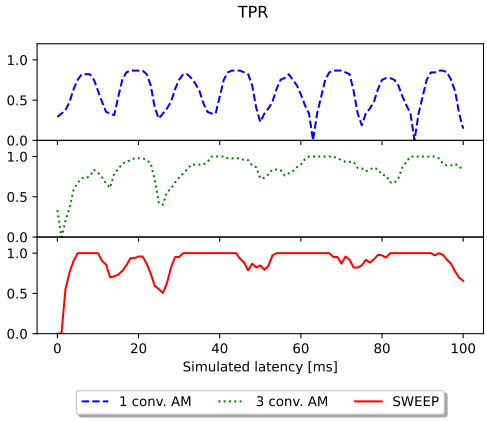


Figure 3. TPR as function of simulated response latency. Each sub-figure represents a different stimulation paradigm. The plotted scores are the grand average of the scores of the individual subject recordings.

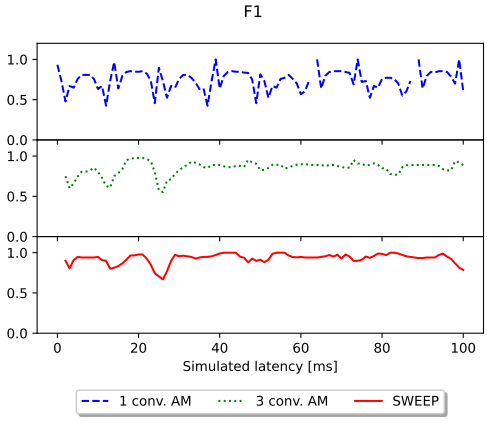


Figure 4. F1 score as function of simulated response latency. Each sub-figure represents a different stimulation paradigm. The plotted scores are the grand average of the scores of the individual subject recordings.

Similar to the accuracy validation with determined response latency, the determined response amplitude can also be used. Here, the determined response amplitude at modulation frequency $f_m = 40Hz$ for a subset of the simulated latency conditions is used. As seen in Fig. 5, the identifiability varies greatly with simulated response latency. We therefore only validated hybrid data with simulated response latencies between $35ms$ to $55ms$, as these represent the latency range of interest for real EEG recordings with modulation frequencies around $40Hz$ (Gransier et al. (2016), Gransier et al. (2017), Schott et al. (2023)).

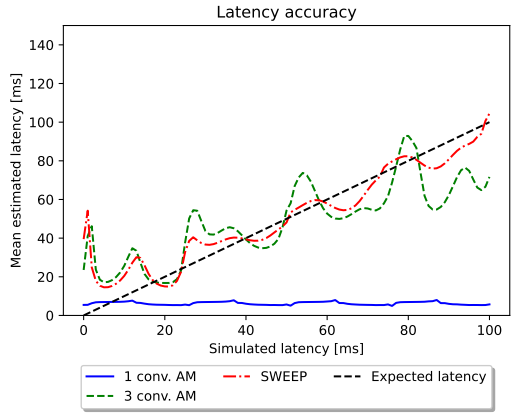


Figure 5. Estimated response latency as function of simulated response latency. Each plot-line represents a different stimulation paradigm. The plotted scores are the grand average of the scores of the individual subject recordings. Only results from hybrid data with SNRs above 6dB are included in the calculation of the mean latency, as responses with lower SNRs are not detectable and estimated latencies orginate from the noise.

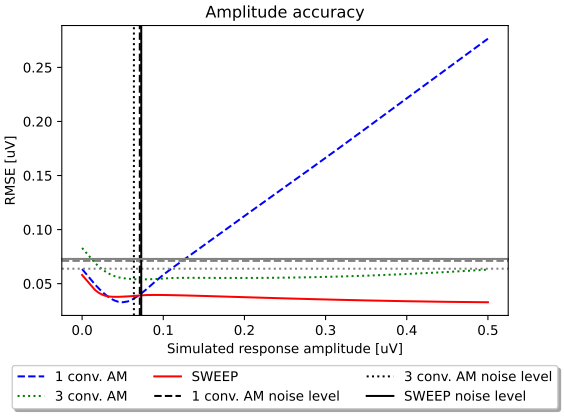


Figure 6. RMSE between estimated amplitude and simulated response amplitude at $40Hz$ as function of simulated response amplitude. Each plot-line represents a different stimulation paradigm. The plotted scores are the grand average of the scores of the individual subject recordings. Only a subset of simulated latency conditions between $35ms - 55ms$ are used.

In Fig. 6 the root-mean-square error (RMSE) between detected and simulated response amplitude for the range of simulated latencies is plotted for each stimulation paradigm. Additionally, the recording noise levels are plotted in gray and black with different linestyles for the different stimulation paradigms. The following observations can be made from the figure.
Using one recording of conventional AM stimulation results quite inaccurate amplitude estimations, as seen from the relatively high RMSE. When using three recordings of conventional AM stimulation, the RMSE significantly drops to a level that is very close to the noise level of the EEG recordings (indicated in grey horizontal and black vertical dotted lines). When using the SWEEP stimulation recording, the RMSE is the lowest and independent of the simulated response amplitude.

3.2. EEG data results

3.2.1. Response latency The determined response latencies of the conventional AM, SWEEP and Shuffled-SWEEP stimulation recordings of all subjects are shown in Fig. 7 ((a)-(d)). The expected latency for the used modulation frequency range lies between $35ms - 55ms$ ((Gransier et al. (2016), Gransier et al. (2017), Schott et al. (2023))), as indicated by the red dashed line. The results from each stimulation paradigm include all analyzed EEG recording channels. The recording sites of the channels are indicated by color, where purple ● stands for contralateral, orange ● for ipsilateral and turquoise ● for central channels. The exact channel is indicated by the marker, which can be found in the legend.

For all subjects and all stimulation paradigms, all determined responses are statistically significant (as determined from HT2), meaning that the determined latencies originate from the elicited response and not from background noise. The response latencies are all within or slightly outside the expected range. Due to the small number of samples, no statistical test is performed to test for a difference between stimulation paradigms. Based on visual inspection, responses detected in recordings with Shuffled-SWEEP stimulation appear to have lower response latencies.

The phase delays of the individual modulation frequencies are shown in Fig. 8. For each subject, one sub-figure is presented ((a)-(d)). The phase delays are offset-corrected so that for each stimulation paradigm, the phase delay at $f_m = 43Hz$ is $-200°$.

3.2.2. Response amplitude The response amplitudes of the individual modulation frequencies are shown in Fig. 9. Again, each sub-figure ((a)-(d)) contains the response amplitudes of the responses elicited by the three stimulation paradigms of the same subject. Based on visual inspection, the response amplitudes are generally consistent across different stimulation paradigms within subjects, except for the ipsilateral recordings of the sweep paradigm in subject 2.

4. Discussion

Recent advances in CI stimulation artifact removal have enabled the use of EASSRs, evoked with clinically relevant stimulation parameters, as neural responses in CI users (Schott et al. (2023)). However, using EASSRs can still be impractical for given reasons.

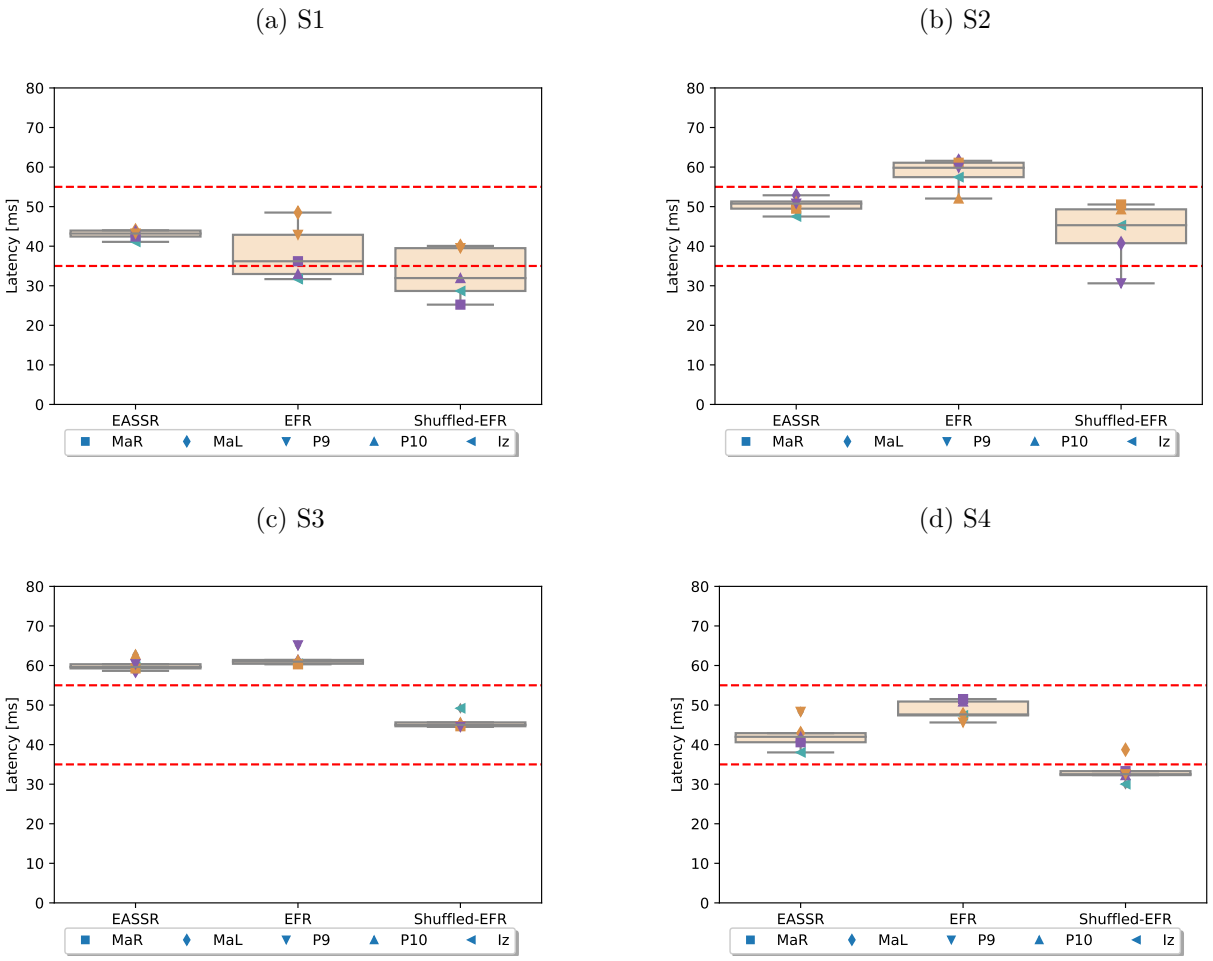


Figure 7. The determined response latencies following stimulation of the different paradigms for all subjects. The expected latency range between $35ms - 55ms$ is indicated with dashed lines. The recording channel name is indicated by the marker, and the location with respect to the CI is indicated by the color (contralateral: ●, ipsilateral: ●, and central: ●).

In this paper, an improved SWEEP stimulation paradigm has been proposed, which is assumed to elicit another EFR. The SWEEP stimulation has several advantages over conventional AM stimulation as used to elicit EASSRs, and it has been shown that the SWEEP stimulation can be further adapted to create more complex stimulation envelopes, closer related to speech envelopes, as seen in Gransier & Wouters (2021), referred to as Shuffled-SWEEP in this paper. EEG data from four CI users has been recorded at high stimulation levels for conventional AM, SWEEP and Shuffled-SWEEP stimulation, and the detected responses have been compared. Furthermore, hybrid EEG datasets have been created by combining artifact-only EEG recordings with simulated responses, to show the benefits of the proposed SWEEP stimulation on artifact removal and response detection with SI.

4.1. Hybrid EEG data

The validation metrics on the hybrid EEG data show that using more input-output information, i.e., more modulation frequencies in the envelope of the stimulation, is beneficial for the SI and hence response detection. Using the SWEEP stimulation with 5 modulation frequencies provides the most accurate results in response detection. Although using three recordings of conventional AM stimulation is good enough to separate artifact and response (as seen from the determined latencies in Fig. 5) and achieve very good detection accuracies as well (Fig. 3), when looking at the TPR as a function of the simulated response SNR (Fig. 10) the following can be concluded. In a recording using SWEEP stimulation, 100% detection accuracy is achieved at simulated response SNRs approximately 2.6 dB lower than those required when using three recordings with conventional AM stimulation. This is due to better response and artifact separation, as seen in the superior latency (Fig. 5) or

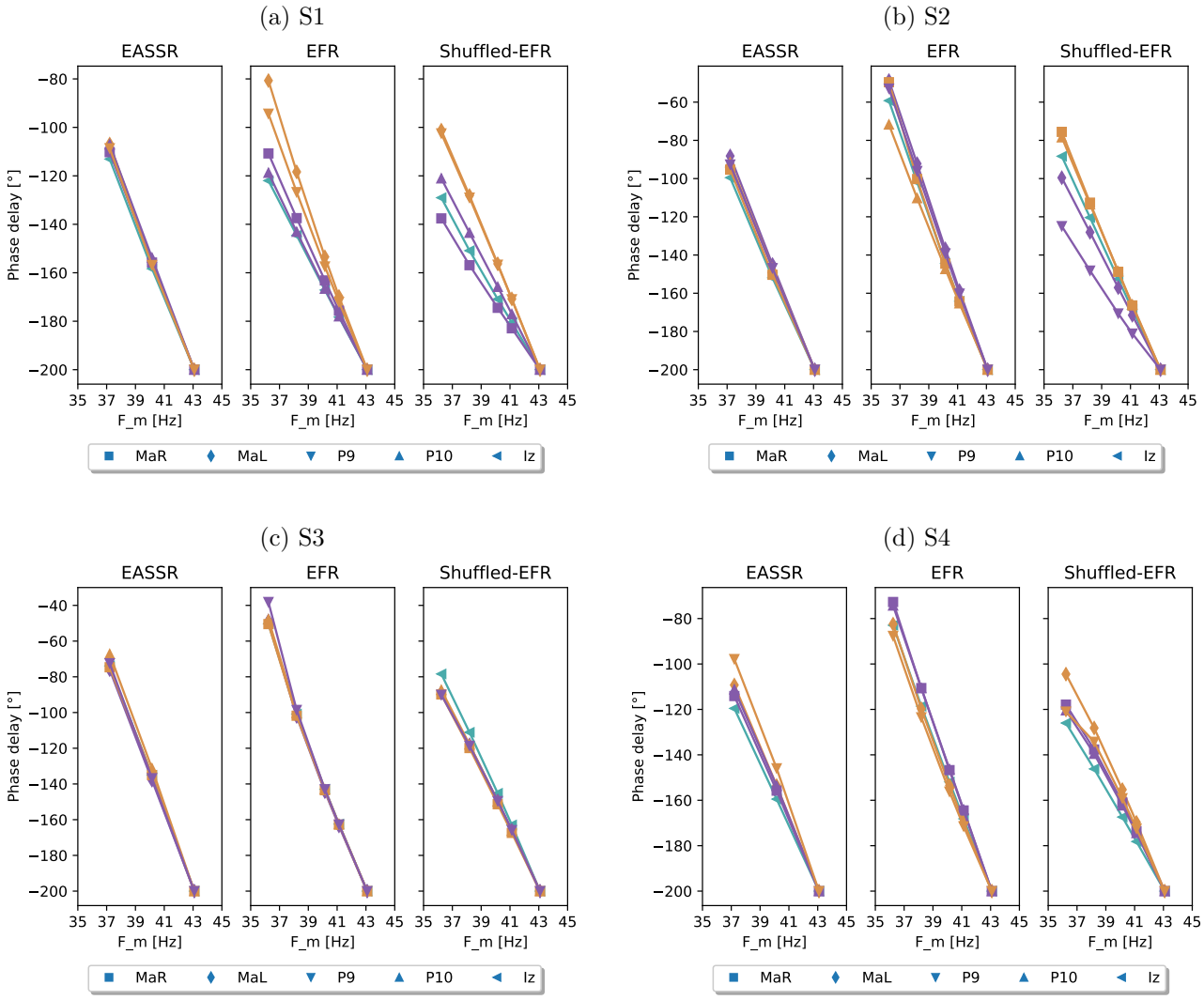


Figure 8. The determined response phases for all four subjects ((a) - (d)). The phase delays are offset-corrected so that for each stimulation paradigm, the phase delay at $f_m = 43Hz$ is $-200°$. The recording channel name is indicated by the marker, and the location with respect to the CI is indicated by the color (contralateral: ●, ipsilateral: ●, and central: ●)

amplitude detection (Fig. 6), which consequently does not lead to false removal of the response when removing the estimated artifact part from the recording. This leads to an 'amplitude-boost' in the response, which is why at lower SNRs the response is better detected. Using the SWEEP stimulation may therefore provide better sensitivity in detecting neural responses in clinical recordings with stimulation levels close to the threshold level of the CI user. Furthermore, under the assumption that the background noise is Gaussian, the SNR is expected to increase with factor of the square root of the number of epochs during epoch averaging (Dobie & Wilson (1996)). An SNR difference of 2.6dB between SWEEP and conventional AM stimulation corresponds to a factor of 1.35, meaning that 1.82 times more averages are needed to achieve equivalent results with three conventional AM stimulation recordings. Consequently, using a SWEEP stimulation with 5 $f_m s$ could reduce the clinical recording time by factor 1.82.

4.2. EEG data

The determined response latencies as seen in Fig. 7 show that for all three stimulation paradigms, neural responses were successfully elicited and detected in the EEG recordings of all subjects and at every recording channel position. Based on visual inspection, Shuffled-SWEEP stimulation appears to elicit neural responses with lower response latencies, which may have two origins.

One, the stimuli might activate different neural generators, leading to higher latencies when more cortical generators are activated and lower latencies when more brainstem generators are activated (Herdman et al.

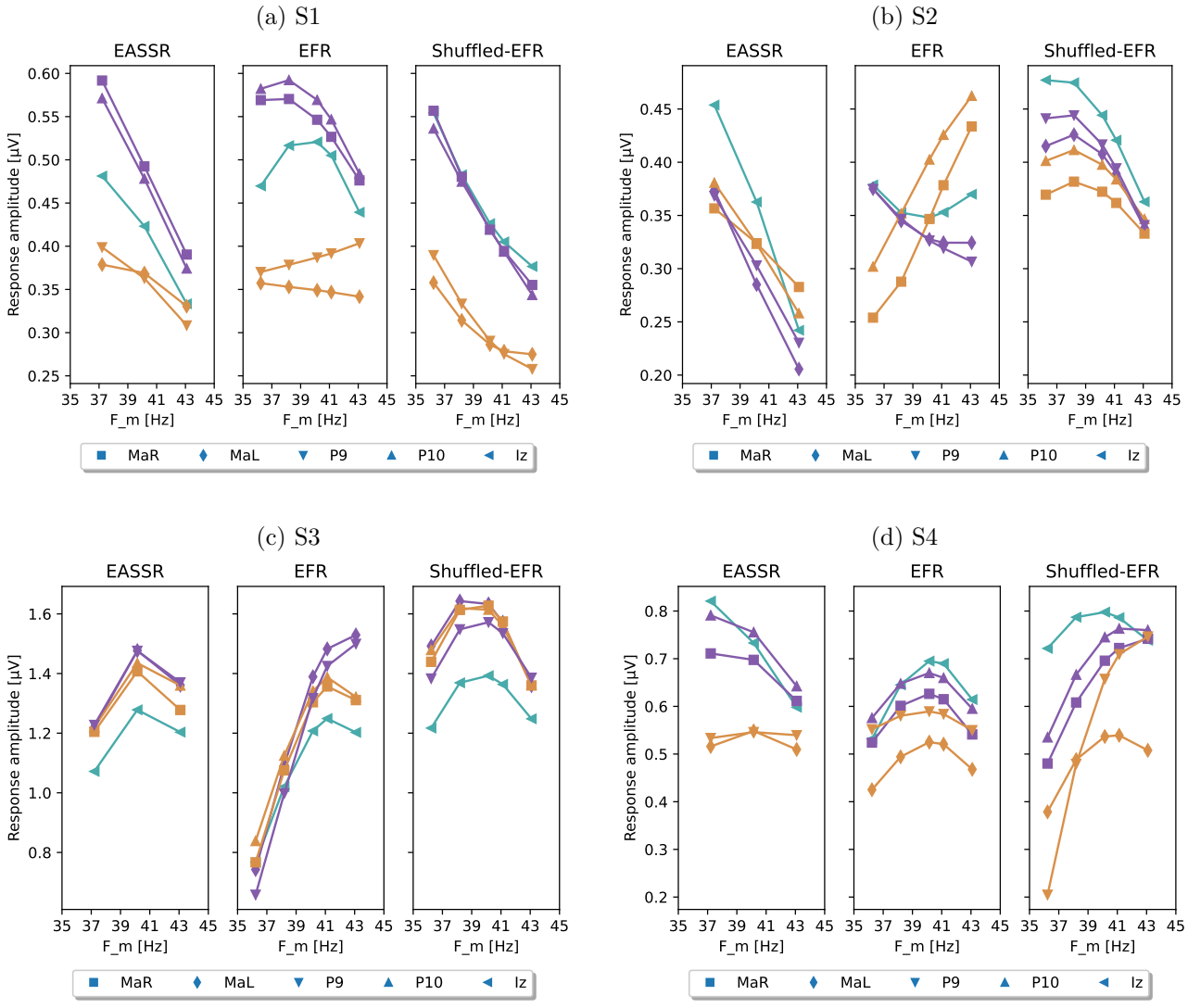


Figure 9. The determined response amplitudes for all four subjects ((a) - (d))

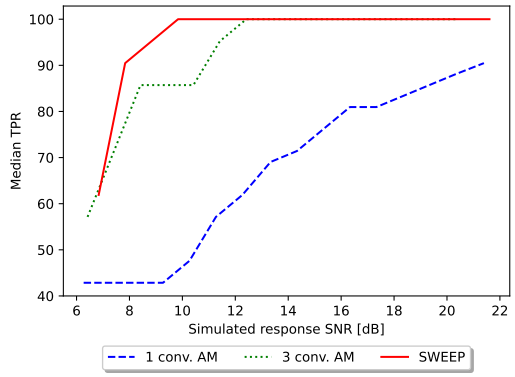


Figure 10. True positive rate (TPR) of the different stimulation paradigms over simulated response amplitudes. The results for simulated response latencies in the range of interest (35-55ms) are averaged.

(2002), Gransier et al. (2016)).

Two, as seen from the validation with hybrid EEG data, a difference in SI accuracy impacts the response detection and response latency estimation. The lower response latencies of the Shuffled-EFRs may therefore originate from the analysis of the EEG data. To gain insight into this possibility, we also used LI as one artifact removal method on the contralateral and central recording channels. These channels are known to be less contaminated by the stimulation artifact (Deprez, Gransier, Hofmann, Wieringen, Wouters & Moonen (2017)), hence LI can indeed successfully be applied. Note that this is only possible with recordings from contralateral channels and with the extremely high sampling rate as used here and seen in Gransier, Carlyon & Wouters (2020) and Schott et al. (2023). With LI, the artifact removal does not depend on the SI accuracy, i.e. is independent of the complexity of the

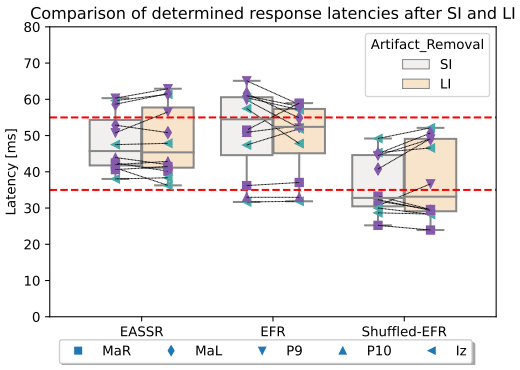


Figure 11. Comparison of determined response latencies after SI and LI. Only EEG recordings from contralateral and central recording channels are compared. No apparent differences in determined response latencies between artifact removal methods indicate a possible difference in active neural generators for Shuffled-EFRs.

envelope of the stimulation. If the determined response latencies of the different recordings are equal after artifact removal with LI, the differences in Fig. 7 may be attributed to differences in SI rather than to differences in elicited neural responses. However, as can be seen in Fig. 11, similar latency results are determined with LI. Again, Shuffled-EFRs appear to have lower response latencies. Furthermore, when comparing response latencies determined after LI with response latencies determined after SI, no obvious difference is observable. Together, this may indicate that the lower latencies in the Shuffled-EFRs effectively originate from a difference in elicited neural response, i.e., a difference in active neural generators, and not from a difference in artifact removal quality. However, due to the small sample size (only 4 subjects), these findings may not be conclusive. Only an investigation on a larger scale can give a clear insight into the differences in elicited EFRs.

5. Conclusion

In this work it has been shown that using a step-wise sweep over 5 modulation frequencies for an amplitude modulated CI pulse train can elicit EFRs that are equal to EASSRs following conventional AM stimulation. Using a hybrid EEG dataset, it has been shown that the proposed SWEEP stimulation improves the separation of response and stimulation artifact, which leads to an 'SNR-boost', that effectively increases detection sensitivity and reduces the required EEG recording time that is needed to detect a response above the noise floor. This greatly benefits the quality of objective fitting in CI users as the accuracy of the T-level detection increases while clinical time decreases. Furthermore, the neural response detection can be performed in an online fashion as the response latency, which is here used to differentiate between response and artifact, can be determined within the same recording, allowing for real-time assessment during ongoing acquisition. This way an early stopping criterion can be implemented in the clinical procedure, which consequently leads to shorter clinical sessions.
In addition to the SWEEP stimulation, a stimulation with 'random' modulating signals (Shuffled-SWEEP) has been proposed and applied in four CI users. Again, neural responses have been elicited and successfully detected. While neural tracking with speech stimuli has been conducted in CI users (Somers et al. (2019), Nogueira & Dolhopiatenko (2022)), there are few results available using a forward modeling approach. Implementing a forward model, as proposed in section 2.3, combined with the proposed stimuli, allows for spatiotemporal assessment of the envelope processing in CI users, as demonstrated by Gransier & Wouters (2021) for normal-hearing individuals. A recent preprint (Federici et al. (2024)) underscores the importance of temporal assessment of speech processing in CI users.

It can be concluded that the combination of the work in Schott et al. (2022), Schott et al. (2023), and here, enables the development of a clinically applicable objective fitting routine, in which EFRs can be used to determine the stimulation levels or assess the temporal envelope processing capability of CI users. Future work should therefore focus on two things. One, investigating the potential difference in neural responses following the different stimulation paradigms, i.e., increasing the sample size and investigating the possible implications of differences in elicited neural responses. Two, the actual development and validation of an objective fitting routine with EFRs.

References

Alemi, R., Nozaradan, S. & Lehmann, A. (2021), 'Free-field cortical steady-state evoked potentials in cochlear implant users', *Brain Topography* .

Bernardi, G., Waterschoot, T. V., Moonen, M., Wouters, J., Hillbratt, M. & Verhaert, N. (2017), 'Measurement and analysis of feedback and nonlinearities for the codacs direct acoustic cochlear implant', *IEEE Access* **5**, 8702–8713.

Brown, C. J., Hughes, M. L., Luk, B., Abbas, P. J., Wolaver, A. & Gervais, J. (2000), 'The relationship between eap and eabr thresholds and levels used to program the nucleus 24 speech processor: Data from adults', *Ear and Hearing* **21**.

Caner, G., Olgun, L., Gültekin, G. & Balaban, M. (2007), 'Optimizing fitting in children using

objective measures such as neural response imaging and electrically evoked stapedius reflex threshold', *Otology and Neurotology* **28**.

Carter, L., Dillon, H., Seymour, J., Seeto, M. & Dun, B. V. (2013), 'Cortical auditory-evoked potentials (caeps) in adults in response to filtered speech stimuli', *Journal of the American Academy of Audiology* **24**.

David, W., Verwaerde, E., Gransier, R. & Wouters, J. (2023), 'Effects of analysis window on 40-hz auditory steady-state responses in cochlear implant users', *Hearing Research* **438**, 108882.
URL: *https://doi.org/10.1016/j.heares.2023.108882*

Deprez, H., Gransier, R., Hofmann, M., van Wieringen, A., Wouters, J. & Moonen, M. (2017), 'Characterization of cochlear implant artifacts in electrically evoked auditory steady-state responses', *Biomedical Signal Processing and Control* .

Deprez, H., Gransier, R., Hofmann, M., Wieringen, A. V., Wouters, J. & Moonen, M. (2017), 'Template subtraction to remove ci stimulation artifacts in auditory steady-state responses in ci subjects', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* .

Deprez, H., Gransier, R., Hofmann, M., Wieringen, A. V., Wouters, J. & Moonen, M. (2018), 'Independent component analysis for cochlear implant artifacts attenuation from electrically evoked auditory steady-state response measurements', *Journal of Neural Engineering* .

Dijk, B. V., Botros, A. M., Battmer, R. D., Begall, K., Dillier, N., Hey, M., Lai, W. K., Lenarz, T., Laszig, R., Morsnowski, A., Müller-Deile, J., Psarros, C., Shallop, J., Weber, B., Wesarg, T., Zarowski, A. & Offeciers, E. (2007), 'Clinical results of autonrt,™ a completely automatic ecap recording system for cochlear implants', *Ear and Hearing* **28**.

Dillier, N., Lai, W. K., Almqvist, B., Frohne, C., Müller-Deile, J., Stecker, M. & Wallenberg, E. V. (2002), 'Measurement of the electrically evoked compound action potential via a neural response telemetry system', *Annals of Otology, Rhinology and Laryngology* **111**.

Dobie, R. A. & Wilson, M. J. (1996), 'A comparison of t test, f test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids', *The Journal of the Acoustical Society of America* .

Drullman, R., Festen, J. M. & Plomp, R. (1994), 'Effect of reducing slow temporal modulations on speech reception', *Journal of the Acoustical Society of America* .

Eeckhoutte, M. V., Wouters, J. & Francart, T. (2016), 'Auditory steady-state responses as neural correlates of loudness growth', *Hearing Research* .

Farina, A. (2000), 'Simultaneous measurement of impulse response and distortion with a swept-sine technique', *Proc. AES 108th conv, Paris, France* .

Federici, A., Fantoni, M., Pavani, F., Handjaras, G., Bednaya, E., Martinelli, A., Berto, M., Ricciardi, E., Nava, E., Orzan, E., Bianchi, B. & Bottari, D. (2024), 'Resilience and vulnerability of neural speech tracking in children with cochlear implants', *bioRxiv* .
URL: *https://www.biorxiv.org/content/early/2024/09/10/2024.02.*

Gordon, K. A., Papsin, B. C. & Harrison, R. V. (2004), 'Toward a battery of behavioral and objective measures to achieve optimal cochlear implant stimulation levels in children.', *Ear and hearing* **25**, 447–463.

Gransier, R., Carlyon, R. P. & Wouters, J. (2020), 'Electrophysiological assessment of temporal envelope processing in cochlear implant users', *Scientific Reports* **10**, 15406.
URL: *https://doi.org/10.1038/s41598-020-72235-9*

Gransier, R., Deprez, H., Hofmann, M., Moonen, M., van Wieringen, A. & Wouters, J. (2016), 'Auditory steady-state responses in cochlear implant users: Effect of modulation frequency and stimulation artifacts', *Hearing Research* **335**, 149–160.
URL: *http://dx.doi.org/10.1016/j.heares.2016.03.006*

Gransier, R., Luke, R., Wieringen, A. V. & Wouters, J. (2020), 'Neural modulation transmission is a marker for speech perception in noise in cochlear implant users', *Ear and Hearing* pp. 591–602.

Gransier, R., van Wieringen, A. & Wouters, J. (2017), 'Binaural interaction effects of 30-50 hz auditory steady state responses', *Ear and hearing* .

Gransier, R. & Wouters, J. (2021), 'Neural auditory processing of parameterized speech envelopes'.

Herdman, A. T., Lins, O., Roon, P. V., Stapells, D. R., Scherg, M. & Picton, T. W. (2002), 'Intracerebral sources of human auditory steady-state responses', *Brain Topography* **15**.

Hofmann, M. & Wouters, J. (2010), 'Electrically evoked auditory steady state responses in cochlear implant users', *JARO - Journal of the Association for Research in Otolaryngology* .

Hofmann, M. & Wouters, J. (2012), 'Improved electrically evoked auditory steady-state response thresholds in humans', *JARO - Journal of the Association for Research in Otolaryngology* .

Hu, H. & Ewert, S. D. (2021), 'Exploring artifact rejection for high-pulse rate electrically evoked

auditory steady state responses in cochlear implants users', pp. 1311–1316.

Hughes, M. L., Brown, C. J., Abbas, P. J., Wolaver, A. A. & Gervais, J. P. (2000), 'Comparison of eap thresholds with map levels in the nucleus 24 cochlear implant: Data from children', *Ear and Hearing* **21**.

Ljung, L. (1999), *System identification : theory for the user*, Prentice Hall information and system sciences series, 2nd ed. edn, Prentice Hall/PTR, Upper Saddle River.

Luke, R., Deun, L. V., Hofmann, M., Wieringen, A. V. & Wouters, J. (2015), 'Assessing temporal modulation sensitivity using electrically evoked auditory steady state responses', *Hearing Research* **324**, 37–45.

Luke, R. & Wouters, J. (2017), 'Kalman filter based estimation of auditory steady state response parameters', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* .

Mao, D., Innes-Brown, H., Petoe, M. A., Wong, Y. T. & McKay, C. M. (2019), 'Fully objective hearing threshold estimation in cochlear implant users using phase-locking value growth functions', *Hearing Research* .

McKay, C. M., Chandan, K., Akhoun, I., Siciliano, C. & Kluk, K. (2013), 'Can ecap measures be used for totally objective programming of cochlear implants?', *JARO - Journal of the Association for Research in Otolaryngology* .

McKay, C. M. & Smale, N. (2017), 'The relation between ecap measurements and the effect of rate on behavioral thresholds in cochlear implant users', *Hearing Research* **346**, 62–70.
URL: *https://www.sciencedirect.com/science/article/pii/S037859551630507X*

Miller, C. A., Brown, C. J., Abbas, P. J. & Chi, S. L. (2008), 'The clinical application of potentials evoked from the peripheral auditory system', *Hearing Research* **242**.

Müller, S. & Massarani, P. (2001), 'Transfer-function measurement with sweeps', *AES: Journal of the Audio Engineering Society* .

Nelles, O. (2001), *Nonlinear System Identification*.

Nogueira, W. & Dolhopiatenko, H. (2022), 'Predicting speech intelligibility from a selective attention decoding paradigm in cochlear implant users', *Journal of Neural Engineering* **19**, 26037.
URL: *https://dx.doi.org/10.1088/1741-2552/ac599f*

Petley, L., Blankenship, C., Hunter, L. L., Stewart, H. J., Lin, L. & Moorea, D. R. (2024), 'Amplitude modulation perception and cortical evoked potentials in children with listening difficulties and their typically developing peers', *Journal of Speech, Language, and Hearing Research* **67**.

Picton, T. W., John, M. S., Dimitrijevic, A. & Purcell, D. (2003), 'Human auditory steady-state responses', *International Journal of Audiology* **42**, 177–219.

Purcell, D. W. & John, M. S. (2010), 'Evaluating the modulation transfer function of auditory steady state responses in the 65 hz to 120 hz range', *Ear and Hearing* **31**, 667–678.

Purcell, D. W., John, S. M., Schneider, B. A. & Picton, T. W. (2004), 'Human temporal auditory acuity as assessed by envelope following responses', *The Journal of the Acoustical Society of America* **116**, 3581–3593.

Schott, J., Gransier, R., Wouters, J. & Moonen, M. (2022), 'Electrically evoked auditory steady state response detection in cochlear implant recipients using a system identification approach', *European Signal Processing Conference* **2022-Augus**, 1243–1247.

Schott, J., Gransier, R., Wouters, J. & Moonen, M. (2023), 'Detection of electrically evoked auditory steady-state responses in cochlear implant recipients with a system identification based method', *Unpublished - in revision* .

Seyle, K. & Brown, C. J. (2002), 'Speech perception using maps based on neural response telemetry measures', *Ear and Hearing* **23**.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J. & Ekelid, M. (1995), 'Speech recognition with primarily temporal cues', *Science* .

Smoorenburg, G. F., Willeboer, C. & Dijk, J. E. V. (2002), 'Speech perception in nucleus ci24m cochlear implant users with processor settings based on electrically evoked compound action potential thresholds', *Audiology and Neuro-Otology* **7**.

Somers, B., Verschueren, E. & Francart, T. (2019), 'Neural tracking of the speech envelope in cochlear implant users', *Journal of Neural Engineering* .

Thai-Van, H., Truy, E., Charasse, B., Boutitie, F., Chanal, J. M., Cochard, N., Piron, J. P., Ribas, S., Deguine, O., Fraysse, B., Mondain, M., Uziel, A. & Collet, L. (2004), 'Modeling the relationship between psychophysical perception and electrically evoked compound action potential threshold in young cochlear implant recipients: Clinical implications for implant fitting', *Clinical Neurophysiology* **115**.

Távora-Vieira, D., Wedekind, A., Ffoulkes, E., Voola, M. & Marino, R. (2022), 'Cortical auditory evoked potential in cochlear implant users: An objective method to improve speech perception', *PLoS ONE* .

Távora-Vieira, D., Wedekind, A., Marino, R., Purdy, S. C. & Rajan, G. P. (2018), 'Using aided cortical

assessment as an objective tool to evaluate cochlear implant fitting in users with single-sided deafness', *PLoS ONE* **13**.

Undurraga, J. A., Yper, L. V., Bance, M., McAlpine, D. & Vickers, D. (2021), 'Characterizing cochlear implant artefact removal from eeg recordings using a real human model', *MethodsX* **8**, 101369.
URL: *https://doi.org/10.1016/j.mex.2021.101369*

Verwaerde, E. et al. (2024), 'In preparation'.

Visram, A. S., Innes-Brown, H., El-Deredy, W. & McKay, C. M. (2015), 'Cortical auditory evoked potentials as an objective measure of behavioral thresholds in cochlear implant users', *Hearing Research* .

Wouters, J., McDermott, H. J. & Francart, T. (2015), 'Sound coding in cochlear implants: From electric pulses to hearing', *IEEE Signal Processing Magazine* **32**, 67–80.