This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2023.3322348

InterlACE Sound Coding for Unilateral and Bilateral Cochlear Implants

Dietmar Michael Wohlbauer, *Member, IEEE*, Wai Kong Lai, and Norbert Dillier, *Life Member, IEEE*

***Abstract*— Objective: Cochlear implant signal processing strategies define the rules of how acoustic signals are converted into electrical stimulation patterns. Technological and anatomical limitations, however, impose constraints on the signal transmission and the accurate excitation of the auditory nerve. Acoustic signals are degraded throughout cochlear implant processing, and electrical signal interactions at the electrode-neuron interface constrain spectral and temporal precision. In this work, we propose a novel InterlACE signal processing strategy to counteract the occurring limitations. Methods: By replacing the maxima selection of the Advanced Combination Encoder strategy with a method that defines spatially and temporally alternating channels, InterlACE can compensate for discarded signal content of the conventional processing. The strategy can be extended bilaterally by introducing synchronized timing and channel selection. InterlACE was explored unilaterally and bilaterally by assessing speech intelligibility and spectral resolution. Five experienced bilaterally implanted cochlear implant recipients participated in the Oldenburg Sentence Recognition Test in background noise and the spectral ripple discrimination task. Results: The introduced alternating channel selection methodology shows promising outcomes for speech intelligibility but could not indicate better spectral ripple discrimination. Conclusion: InterlACE processing positively affects speech intelligibility, increases available unilateral and bilateral signal content, and may potentially counteract signal interactions at the electrode-neuron interface. Significance: This work shows how cochlear implant channel selection can be modified and extended bilaterally. The clinical impact of the modifications needs to be explored with a larger sample size.**

***Index Terms*— Channel interaction, cochlear implant, signal processing, speech intelligibility, spectral resolution.**

Manuscript received 19 April 2023; revised 22 August 2023; accepted 28 September 2023. The work leading to this deliverable and the results described therein has received funding in part from the people Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013 under REA grant agreement number PITN-GA-2012-317521 and in part by the University of Zurich and the Truus und Gerrit van Riemsdijk Foundation for the completion of the PhD thesis. *(Corresponding author: Dietmar Michael Wohlbauer.)*

Dietmar Michael Wohlbauer was with the Institute of Neuroinformatics (INI), University of Zurich (UZH), Zurich, Switzerland, and the Laboratory of Experimental Audiology (LEA), Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital Zurich (USZ), 8006 Zurich, Switzerland. He is now with the Harvard Medical School (HMS), Boston, MA USA, and the Emerging Auditory Research (EAR) Lab, Department of Otolaryngology, Massachusetts Eye and Ear (MEE), MA 02114 USA (e-mail: dm.wohlbauer@gmail.com).

W. K. Lai is with NextSense, formerly Sydney Cochlear Implant Centre (RIDBC), Sydney, Australia.

N. Dillier is with the Laboratory of Experimental Audiology (LEA), Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital Zurich (USZ), and the University of Zurich (UZH), Zurich, Switzerland.

I. Introduction

Although the performance of clinically available cochlear implant (CI) sound processing methods is sufficient for many everyday listening situations, present-day CI coding strategies often distort important signal parameters and, therefore, impose limits on the amount of acoustic spectral information conveyed electrically to the available nerve fibers [1]–[4].

One way of assessing the CI users' performance is the so called speech reception threshold (SRT), the recognition threshold for correctly understood words. Another method to assess the performance is the spectral discrimination threshold, i.e. the spectral resolution. These measures can be used to evaluate the effect of CI parameter variations such as the number of active implant channels, split signal content, and unilateral versus bilaterally fitted CIs, and thus to investigate the efficacy of coding strategies and their behavior in different listening situations.

Coding strategies such as the Advanced Combination Encoder (ACE) and Continuous Interleaved Sampling (CIS) show good performance in speech intelligibility and spectral discrimination experiments [5]–[7]. Compared to the ability of normal-hearing (NH) subjects in such experiments the results are worse for severe or profound hearing-impaired CI recipients. Comparative studies showed significantly reduced performance for speech recognition and spectral resolution through ripple discrimination assessments [6], [8], [9].

The impact of the number of actively stimulating electrodes on the recognition of real and synthetic vowels, consonants, monosyllabic words, and sentences in quiet has also been investigated before [10], [11]. Higher channel numbers had a positive effect on speech intelligibility outcomes, but an increase beyond more than nine active channels showed no further improvements. By adding noise to the listening situations, later studies reached the same conclusion of no additional perceptual benefit above a certain number of active channels [12], [13]. A more recent study showed improvements for numbers above nine until the maximum available number of channels, indicating the importance for the transmission of more complex signals [14].

ACE coding restricts the number of available active electrodes per stimulation frame preserving only strong signal content while discarding low-intensity content, thereby limiting its information transmission capacity but also channel interaction at the electrode-neuron interface. CIS coding, on the other hand, can transmit signals on all available electrodes sequentially, but cannot completely avoid stimulus interactions of the processed and transmitted signal. Therefore, increasing the number of active channels may lead to more spectral content available after the processing but leads to an increased electric crosstalk at the neuronal interface as consequence of the increased number of activated electrodes. This crosstalk cannot simply be reduced due to anatomical factors such as the actual size of the cochlea and the implanted electrode array or the electric conductivity of the perilymph, endolymph, blood, surrounding tissue, etc.

Furthermore, clinical coding strategies are unilateral signal processing approaches and provide no synchronization or communication between bilateral processors. The possibilities for bilateral extensions increased over the last years and are explored more intensely to show the benefits of bilateral hearing in more complex listening situations such as speech-in-noise or spatial perception. Assessments of CI coding strategies in a bilateral setup revealed significant performance improvements in numerous studies when compared to unilateral processing [7], [15], [16]. An experimental bilateral implementation of ACE coding, addressed as binaural N-of-M (BINOM), assessed speech intelligibility when CI channels are linked between the bilateral processors [17]. Speech tests in noise at 90 degrees on the horizontal plane showed improvements for linked over unlinked conditions, which was confirmed in a detailed follow-up study [18].

Apart from previously indicated performance improvements, unilateral and bilateral coding approaches are not yet overcoming certain disadvantages of limited electrode numbers and large electric crosstalk. Thus, several studies explored the idea of an interleaved signal content and the reduction of signal interactions. For example, speech performance measured via syllables and vowels in a split-formant condition in the ipsi- and contralateral ear, thus generating a hearing sensation of different sounds in each ear, did not show improvements in comparison to the binaural signal representation for NH listeners [19]. In a follow-up study speech intelligibility was also significantly affected by the dichotically presented formants. An improvement of $2\,\mathrm{dB}$ in diotically and binaurally presented speech in noise was found without any additional improvement of switching the odd and even bands between the left and right ear periodically [20]. The dichotical presentation of vowel-consonant-vowel and consonant-vowel syllables in noise was further explored and significant perceptual improvements found [21].Another study showed improvements of about $5\,\mathrm{dB}$ for spectrally comb split monosyllabic words for NH study subjects [22]. The impact of splitting the signal content on speech intelligibility in quiet with a similar idea was further explored with signals processed by a pair of comb filters and balanced perceptually based on auditory critical bandwidth [23]. Improvements for speech intelligibility suggested a useful application of such signal separating algorithms in future binaural devices.

To increase the spectral discrimination performance, the separation of active channels to two cochlear implants could counteract current interactions among the array electrodes as a limiting factor of poor spectral resolution [24]. This method was explored in more detail, however only for speech intelligibility, but found that interleaved and noninterleaved strategies behave relatively similar [25]. On the contrary, a different study assessed the effect of spectral masking and found that comb filtering signals reduced spectral masking in subjects with sensorineural hearing loss [26].

Even though study outcomes seem to be mixed for binaural dichotic signal presentation, attempts to overcome current limitations, therefore, to fill the gap of performance deficits with new CI coding ideas are more relevant and timely than ever. To successfully incorporate previous findings with the intention to overcome the drawbacks with channel numbers and signal interactions, the present work explores a CI coding paradigm, which is based on perceptual fusion, respectively, two assumptions, (a) the temporal and (b) spectral integration of evoking pulsatile signals. The temporal fusion of information can be understood as the ability of the auditory system in electrically evoked hearing to integrate temporally separated signals. Spectral integration can be understood as the ability to perceptually fuse spectrally separated information. Several studies established that temporal and spectral integration takes place in CI users similarly as in NH subjects [19], [27]. It was shown that signals are temporally integrated with subjective variations in the integration function on different electrodes [27]. An explanation for the variations could be individual differences in loudness functions which unfortunately were not recorded in that study. Spectral fusion or integration of acoustic information into one sound image in the auditory cortex was investigated in NH and bilateral CI users and supported through several studies. Findings showed an at least comparable perception of sound achieved by splitting signals between ears [19]–[26]. On average it can be assumed that CI subjects follow the same behavior in time and frequency as it was observed in NH subjects, except for the overall decreased stimulus intensity increments with increasing delay times.

Given the perceptual ability to fuse temporally and spectrally separated signals and to re-evaluate the benefit of interleaved signal content in severe-to-profound hearing-impaired bilaterally implanted individuals, the current study investigates the novel coding paradigm, denoted as InterlACE, with the hypotheses to improve speech intelligibility and spectral resolution. The findings should promote efforts towards new bilateral CI coding ideas and ultimately improve signal transmission and perception.

II. Methods

All procedures performed in this work were approved by Swissmedic (Reference-Nr. 10000818) and the Zurich ethics committee (BASEC Reference-Nr. 2020-00157) in agreement with the Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. All subjects gave their informed consent.

A. Subjects

A total of five (three females, two males) experienced bilaterally implanted CI users (average age of 47.6 years, range between 20 to 68 years) participated in the study. Before implantation, subjects were deaf for an average of 4.9 years on the left and 3.6 years on the right. Etiologies varied among subjects and were not always clear to define. Some subjects suffered from prelingual hearing loss in one ear, others suffered from progressive post-disease types of hearing loss. All subjects were implanted for an average of 6.2 years on the left and 9.3 years on the right. The main parameters and etiologies are shown in Table I. For the unilateral assessments, participants were asked about their preferred ear in daily listening, indicated in bold. Three of the subjects (CI01, CI02, and CI04) reported better performance and preference for their first, two (CI03 and CI05) for the later implanted ear.

All subjects were considered as good performers in standard clinical adaptive speech intelligibility assessments, the Oldenburg Sentence Recognition Test (OLSA) at a presentation level of 65 dB sound pressure level (SPL), with at least 70% correct answers in quiet for both ears. Nearly all subjects performed above the set threshold with 90% and higher. On average, study participants showed a symmetric performance with 87.2% on their left and 88.2% on their right ear. All subjects were fluent in the German language, bilaterally implanted for at least six months with 22 active electrodes, clinically used the ACE coding strategy, except for one ear, and had the ability to perform speech intelligibility procedures in quiet and noisy environments. In addition, a good health condition was crucial for participation.

As shown in Table II subjects are implanted with Nucleus® implants of the Cochlear® Corporation and implant types CI24RE(CA), CI422, CI512, or CI522. The original maps had standard channel stimulation rates between 500 pps (pulses per second) and 900 pps, and six to eight maxima. The majority of the maps used the ACE coding strategy, with only one exception using MP3000. A pulse width of 25 µs was used in seven of the ten tested ears. Three ears were fitted with 37 µs due to compliance maxima measures of the implanted electrode arrays.

For the evaluation of the InterlACE coding strategy, the individual clinical maps were adapted regarding the stimulation rate, active channel maxima, and the pulse width. In an individual fitting session, the most recent clinical fitted maps for the left and right ear formed the base of the adjusted research maps. Hence, the levels of the individually updated threshold (T-level) and comfort levels (C-level) closely matched the initial clinical maps and were adjusted mostly for C-level per channel. To achieve a maximum pulse rate for the coding of interlaced active channels the initial channel stimulation rate and active channel number were increased. The basis for the InterlACE coding paradigm was set to 1200 pps and eleven maxima with a pulse width of 25 µs. The updated rate and maxima selection were defined based on the processor limits extracted from the Nucleus MATLAB® Toolbox (NMT), a toolbox addressed hereinafter, and the implementation of the ACE coding strategy. The limit for the maximal implant rate is at 14.4 kHz and the new setting was the closest possible optimal approximation with the maximum of 13.2 kHz. Any other combination of channel stimulation rate and number of maxima would have resulted in a suboptimal fitting or a stimulation rate far below or above the possible implant rate.

The eleven maxima are a result of the channel selection method and are essential to cover the whole electrode array for stimulation. C-levels of the subject maps are set to be perceived with the same loudness on all electrodes and both ears. Research maps were fitted carefully following the clinical procedure to facilitate optimal sound perception for the subject and fulfill the requirements of the study. The patients' own speech processors and maps were left unchanged throughout the whole experimental sessions.

All participants could communicate with or without the processor. The unilateral tasks were performed with active contralateral processor but the participants were asked to switch it off if noises in the laboratory or processor noise were affecting the task execution. For bilateral experiments, both main processors had to be exchanged by the experimental processor with the corresponding generated research maps.

B. Hard- and Software

Two types of CI signal processors were used for the study: the Nucleus® CP910, a clinically used sound processor of the Cochlear® Corporation, that served for the fitting of the research maps, and the RF Generator XS, a streaming platform with the ability to transmit the unilaterally and bilaterally processed stimuli, thus a link between the experiment software and the CI subject. The fitting of the study maps was performed with Custom Sound® (CS) Pro, Cochlear's implant fitting software. Experiment signals were generated in a custom interface, denoted as ZORRO, programmed in MATLAB® R2013b. ZORRO was programmed as a flexible tool for psychoacoustic experiments for CI stimulation, and to combine multiple procedures in one tool to ensure consistent

TABLE I: Study subject details

Subject	Gender	Age[a]	Duration of implant use[a]		Deafness Duration[a]		Etiology	
			Left	Right	Left	Right	Left	Right
		(yrs)	(yrs)	(yrs)	(yrs)	(yrs)		
CI01	f	20	5	**8**	∼1	12.9	Con*	Con
CI02	m	68	3.5	**16.5**	∼12	0.6	M.M*	M.M*
CI03	f	48	9	**4**	1.8	0.4	Her	Sens
CI04	m	54	11	**13**	∼6	3.2	Men*	Men*
CI05	f	48	**2.5**	5	∼3.5	1.1	ISSHL	Vacc**
Mean		47.6	6.2	9.3	∼4.9	3.6		

[a] Years in (yrs) * Decrease in residual hearing ** Suspicion
M.M = Morbus Menière Con = Congential Men = Meningitis
Her = Hereditiary Sens = Sensorineural Vacc = Vaccination
ISSHL = Idiopathic sudden sensorineural hearing loss

TABLE II: Implant types and clinical processor settings

Subject	Implant		Stimulation Strategy		Pulse Rate[a]/ Maxima		Pulse Width[b]	
	Left	Right	Left	Right	Left (pps/n)	Right (pps/n)	Left (µs)	Right (µs)
CI01	CI512	CI24RE	ACE	ACE	900/8	900/8	25	37
CI02	CI512	CI24RE	ACE	MP3000	900/8	500/6	25	25
CI03	CI422	CI522	ACE	ACE	900/8	900/8	25	37
CI04	CI512	CI24RE	ACE	ACE	900/8	720/8	25	37
CI05	CI512	CI512	ACE	ACE	900/8	900/8	25	25

[a] Pulses per second and number of maxima in (pps/n)
[b] Pulse width in (µs)

execution of the planned experiments. The use of MATLAB® version R2013b allowed access to the clinical database for retrieval of the experimental maps. Signals were processed via the NMT research platform developed by Cochlear® . NMT includes functions for the conversion of the acoustic signal into electrical stimulation patterns [28]. The patterns can be streamed directly to the CI subjects' implants as pulse sequences via the Nucleus Implant Communicator (NIC). This software enables researchers to generate stimulation sequences of arbitrary stimuli on a computer which is then connected to the appropriate hardware to transmit the sequence of interest as a direct stream to the CI recipient [29].

C. InterlACE Sound Coding

InterlACE coding is based on the processing chain of the ACE coding strategy. In ACE coding, the audio signal from the microphone output of the processor is first pre-emphasized. In the next step, temporal frames of the signal are decomposed into frequency channels with a Fast Fourier Transform (FFT) bandpass filter bank, and the amplitudes of the 22 available frequency bands are extracted. Strongest bands are preserved to stimulate the corresponding electrodes while weak bands are discarded, defined as N-of-M maxima selection. The number of frequency bands rejected is calculated from the maximum number of bands minus the number of selected or active bands. In the final processing stage, the acoustic amplitudes of selected electrodes are mapped and compressed into the subject's dynamic range between measured T-levels and maximum C-levels for the electrical stimulation [30].

In InterlACE coding, the microphone gain functions, the input filter bank with its defined channel frequency cutoffs, and the envelope extraction processing step are kept similar to the ACE coding. The automatic gain control (AGC) is deactivated due to signal calibration before the experiments. The conventional channel selection step of ACE coding is replaced by an alternating channel selection algorithm, schematically presented in Fig. 1 (top).

In comparison to ACE coding, shown in Fig. 1 (bottom), the pulse separation increases in both space between adjacent electrodes along the array and time between repetitive pulses presented to the same electrode. Hence, associated with the channel selection of InterlACE coding is a change of the channel stimulation rate due to the selection paradigm of even and odd channels per processing frame. Consecutive pulses on the same channel are separated by at least one frame which leads to half the channel stimulation rate. Initially, the rate is defined with 1200 pps which leads to an interlaced rate of 600 pps. The effect of various stimulation rates between 400 pps, 800 pps, and the subjects' clinical rates above 1500 pps for consonants and speech perception was investigated before with a similar mean performance in favor of the subjects' known stimulation rates [31]. Also, for the recognition of vowels, consonants, consonant-nucleus-consonant (CNC) words, and IEEE sentences, changes in target stimulation rates between 600 pps and 4800 pps led to similar conclusions [32]. A comparison of the influences of rate changes between 500 pps and 1200 pps on speech

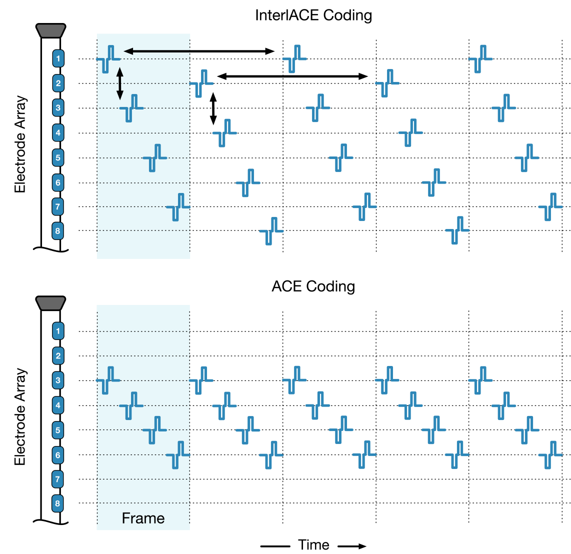

Fig. 1: Schematic channel activation of InterlACE (top) and reference ACE coding (bottom) with four maxima, indicated temporal, and spectral separation of active electrodes, i.e. stimulation pulses per electrode

perception in older CI users ($\geqslant$ 65 years of age) resulted in comparable performance but significant individual differences for lower-than-default stimulation rates [33]. The studies argue that the similar performance throughout different stimulation rates may be explained by the limitations in detecting amplitude modulations, differences in loudness growth, or auditory nerve firing behavior. Hence, based on previous evidence, a reduction to 600 pps does not significantly influence the perception of signals such as speech, vowels or consonants [31]–[35].

1) Unilateral Spectrotemporal Interlacing: InterlACE coding can perform sound processing unilaterally as well as bilaterally. Compared to ACE coding, InterlACE thereby transmits the signals on all available electrodes within short consecutive processing frames. To achieve increased signal transduction by preserving omitted content and avoiding unintended electric crosstalk, the N-of-M maxima selection of ACE has to be re-defined.

The unilateral implementation of InterlACE can be seen as either the top or the bottom part of Fig. 2, which represents the block diagram of InterlACE. Instead of searching for the strongest bands, either even or odd N-of-M bands define the active channels that are mapped to the electrodes. The algorithm is set to select eleven active channels per frame, which is defined in the subject map. With the intention to stimulate all 22 available electrodes over time, the channel selection switches from even to odd and vice versa in the consecutive frames. The decision criterion, included in Fig. 2, defines the interlaced channel selection per time interval. In the final processing stage, the channel mapping, the magnitudes or stimulation current levels of the eleven active electrodes are mapped to the corresponding electrodes.

The modifications in the signal processing chain were executed in the NMT research platform. Influencing factors such as the implant rate, channel stimulation rate, and the

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2023.3322348

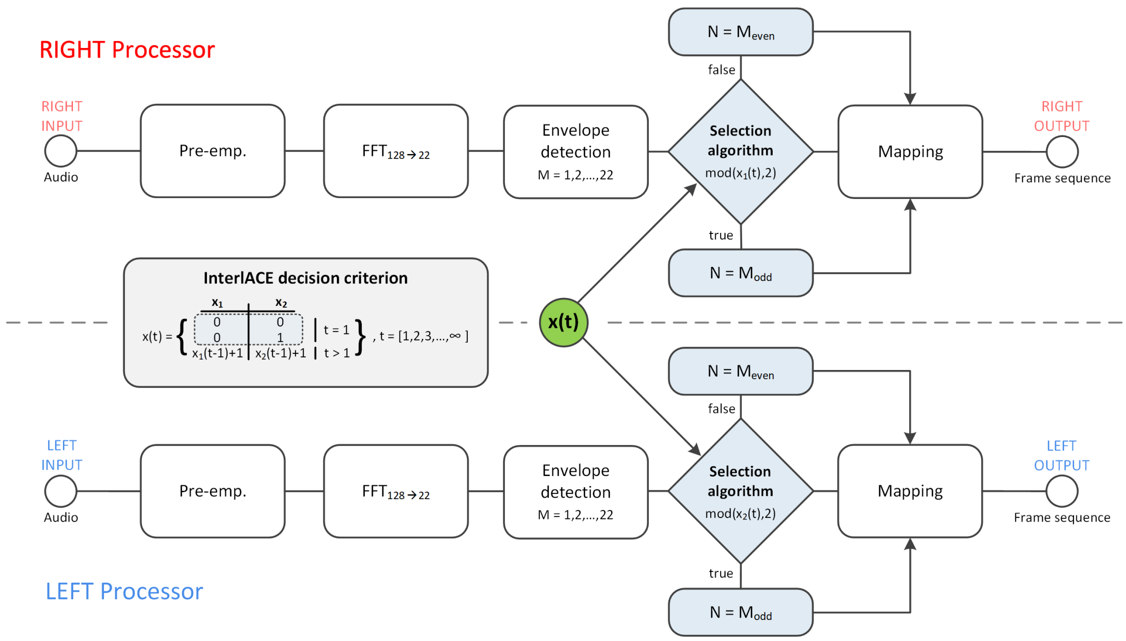

Fig. 2: Block diagram of InterlACE coding. The top or bottom processing chain individually represents the unilateral implementation of InterlACE. The decision criterion defines the link for the bilateral processing and its variations.

number of active electrodes were discussed earlier in this work.

2) Bilateral Extension for Spectrotemporal Interlacing: The algorithm of unilateral interlacing stays unchanged for the bilateral interlaced channel selection, shown in the block diagram of the bilateral InterlACE strategy (Fig. 2). To increase the information compared to the unilateral coding, this method is based on a smart bilateral distribution of the interlaced selected bands. By introducing one important parameter, here defined as $x(t)$ with t standing for the time interval or processing frame, a link between both implants is established. By defining $x(t)$ through the decision criterion, synchronous processing of the typically independent implant processors is triggered.

As described for unilateral InterlACE, the frequency band selection is consecutively chosen to be even or odd. Hence, two variations can be defined for a bilateral case. By setting $x_1(t) = x_2(t)$, the band selection for the left and right processors is the same, presented in Fig. 3 (top). When switched on, the processors will in parallel perform the same interlacing procedure synchronized in their timing. This setting thus defines the first version, addressed as bilateral linked InterlACE coding. A drawback of this selection with the same frequency bands chosen in both processors are the informational gaps, consequently discarding either even or odd frequency bands completely within one interval of t. Information that falls into these gaps is lost. On the other hand, choosing $x_1(t) = x_2(t)$ would introduce informational redundancy throughout the bilateral processing, since the same channels are active on both processors in the same processing frame. This redundancy leads to a waste of available signal resources by ignoring half the information per time interval. To avoid this effect, the signal content can be distributed more effectively in the $x_1(t) \neq x_2(t)$ setup, presented in Fig. 3 (bottom). This processing ensures that either even bands in the left processor and odd bands in the right processor or vice versa are selected. Hence, every single frequency band can bilaterally be represented in a single processing interval. This parameter definition forms the second version of this strategy, addressed as bilateral alternatively linked InterlACE coding. Recapturing the ability of temporal and spectral integration of information by the auditory system, such a method has the potential to preserve the fully available informational content of 22 channels per time interval [19], [23], [25], [27].

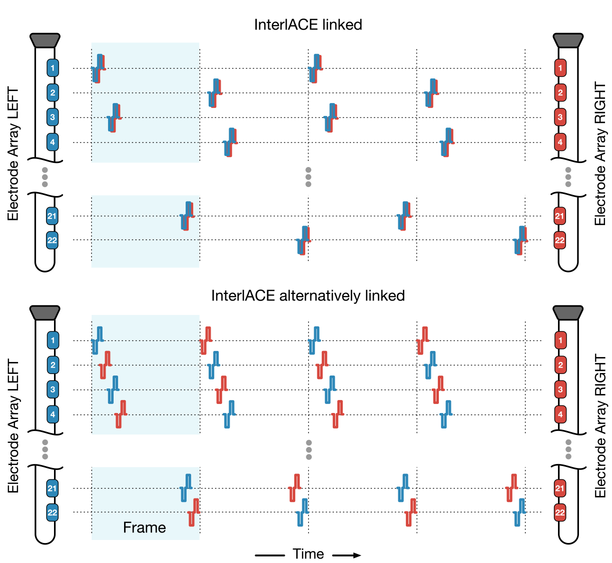

Fig. 3: Schematic channel activation of the linked InterlACE variations, InterlACE bilateral (top) and the InterlACE alternative (bottom), with eleven active electrodes per frame.

3) Stimulation Pattern and Verification: Verification of the InterlACE algorithm was carried out using the output stimulation patterns from both InterlACE and ACE coding and the charge histogram. Fig. 4 shows a comparison of the output patterns, in the form of electrodograms and proportion of

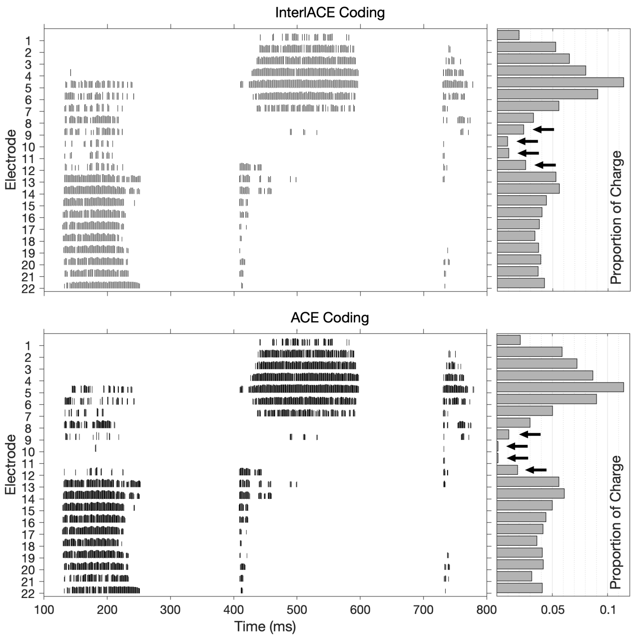

Fig. 4: Electrodograms (left) and proportion of charge (right) for the Freiburg monosyllabic word "Axt" obtained by InterlACE (top) and ACE coding (bottom). Low channel numbers of the electrodogram refer to basal channels and higher frequency content. The proportion of charge is depicted as histogram, compiled per channel between 0 and 1. The arrows indicate the increased channel activity of InterlACE.

charge per channel, for the Freiburg monosyllabic word "Axt" (from the Freiburg monosyllabic word test lists).

The electrodograms contain the stimulus pulses presented on each of the 22 electrodes over time as narrow vertical bars, with the amplitude of each pulse denoted by the height of the bar. Since ACE only allows eleven channels per processing frame, the distribution of active channels is relatively sparse. This can be seen by the empty channels in Fig. 4 (bottom) for the ACE coded signal, especially obvious for the vowel portion around $300\,\mathrm{ms}$ on channels nine to twelve. In InterlACE coding, an increased number of channels is selected within consecutive processing frames, even for signals of low complexity, shown by the example monosyllable, thus resulting in an output pattern including discarded pulses of the ACE coding, shown in Fig. 4 (top). Due to the alternating channel selection algorithm and the reduced channel stimulation rate of InterlACE compared to ACE, the pulses are shifted further apart, hence the electrodogram appears less dense.

The increased channel activity can be verified with the charge histograms, shown as horizontal bar graph to the right of the electrodograms of Fig. 4. The histograms represent the proportion of current delivered per channel compared to the overall delivered current. An increase in charge can be seen for regions with higher channel activity on channels nine to twelve in InterlACE compared to ACE processed stimuli, indicated by arrows, supporting the increased channel activation.

D. Procedures and Stimuli

Speech reception thresholds were measured with the Oldenburg sentence test (OLSA). The OLSA has been developed analogous to the Swedish matrix sentence test for the evaluation of speech intelligibility performance in German [36]–[40]. The list of sentences can be used for assessments in quiet and noise and due to the random word selection, repeated measures can be performed. The sentence matrix is based on lists of name, verb, numeral, adjective, and object and comprises ten available elements per category. The word sequence is generated by a category-wise selection of a single element to form semantically unpredictable sentences as a result of any possible combination.

Training was executed before the experiment, which consisted of one to two lists of 20 sentences in quiet and noise until a constant threshold was reached. The speech intelligibility experiment was executed in background noise and for two lists of 30 sentences per condition. The noise level was at a constant $60\,\mathrm{dB}$ SPL, sentence presentation levels were varied in an adaptive manner dependent on the subject responses, respectively correctly answered words, starting from the initial signal-to-noise ratio (SNR) of $0\,\mathrm{dB}$ after every presented list. The $60\,\mathrm{dB}$ SPL noise level was chosen in order to avoid signal clipping of the acoustic signal in case of large positive SNR values during the assessment, i.e. high signal levels compared to the reference noise level. The adaptive process continued for the whole list of sentences or until the SNR leveled off for at least ten representations in a row. The adaptive procedure was defined to reach a speech reception threshold (SRT) of 50% correct words. The algorithm adaptively selects larger level steps until the fifth presentation and then changes to smaller steps until the end of the sentence list [41]. The average SRT was calculated as the geometric mean from the last ten values for the training lists and the last 20 values in the test condition. Lower or negative SRT outcomes of the experiment indicate a better performance. All sentences for the training in quiet and noise and the experiment in noise were presented in randomized order.

Spectral resolution was assessed with a spectral ripple discrimination task. The spectral discrimination stimuli were generated as logarithmically spaced ripples. Logarithmic spacing approximates the parameters of the auditory system and is therefore thought to correlate with the acoustics of speech [6], [42]. The stimuli frequency range was adjusted to the CI channel band filters extracted from the NMT research platform ranging from $188\,\mathrm{Hz}$ to $7.94\,\mathrm{kHz}$. The raw signal was defined as white noise sampled at $44.1\,\mathrm{kHz}$. Ripple noise stimuli were generated at $60\,\mathrm{dB}$ SPL with an approximate peak-to-valley ratio of $30\,\mathrm{dB}$ and defined to a maximum length of $500\,\mathrm{ms}$. Reference full-wave rectified sinusoidal spectral envelope stimuli were initialized with a zero phase shift and inverted stimuli envelopes were shifted by $\pi/2$. To avoid fine structure cues and signal edge effects, the ripple starting phases were randomized with phase shifts of $\theta = x \times (\pi/8)\ radians$, where $x \in [0, 3]$ to only allow shifts of $\theta < \pi/2\ radians$ corresponding to the inverted ripple phase. Signals were onset/offset gated by a raised-cosine ramp of $50\,\mathrm{ms}$ as rise/fall time, shorter than in previous studies [6], [42], [43]. A silent duration of $100\,\mathrm{ms}$ was added before each stimulus to allow the implant to power-up phase, a specification defined for Cochlear® Nucleus implants. The final ripple envelope frequency was varied in 22 steps: 0.09, 0.18, 0.27, 0.37, 0.46, 0.55, 0.64, 0.74, 0.92, 1.11, 1.29, 1.48, 1.66, 1.88, 2.03, 2.77, 3.88, 5.55, 7.96, 11.11, 15.74, and 22.22 ripples per octave (rpo). Output ripple signals were finally sampled to $16\,\mathrm{kHz}$ to match the signal input requirements for the NMT signal processing.

The spectral ripple discrimination assessment commenced at the easiest condition of 0.09 rpo and followed a three-alternative forced-choice (3AFC), two-down one-up procedure (2D1U) that converged at the 70.7% correct threshold and the resulting spectral resolution [44]. With an interpulse interval of $500\,\mathrm{ms}$, the reference stimuli were presented twice, and the inverse test stimulus once. Each incorrect response decreased the spectral ripple complexity, i.e. less rpo, two correct responses increased the complexity. The step size used in the algorithm followed a binary search method by dividing the possible list length after every step by two until the step size reaches its minimum. This is a fast method to find the convergence level within around four to five steps. Phase-shifted stimuli were randomly selected and level roved by $\pm\,2\,\mathrm{dB}$ to avoid loudness effects. The participating subjects were asked to indicate the differently perceived stimulus by pressing the corresponding button in the user interface. Training was executed similarly to the speech intelligibility assessment once or twice, until a constant threshold was reached. The main assessment was performed twice. No feedback was provided during the procedure. The probe threshold level was estimated after ten reversals for each run as the geometric mean of the ripple frequencies for the last six reversals.

OLSA stimuli from standard clinical speech intelligibility testing, as well as the spectral ripple test stimuli, were initially derived as WAV files. To present these stimuli via direct streaming to the implant at specific loudness levels, the playback levels were calibrated as follows [45]. First, calibrated pure tone signals at specific dB SPL levels were presented to the sound processor inside a hearing aid test box (Interacoustics HIT440), and the processor's RF output captured and decoded the electrical signals by DIET hardware and RFcap software [46]. These provided the reference values for the direct streaming signals. Next, MATLAB®generated sinusoids were presented via a sound card to the direct input of the sound processor, and their amplitudes adjusted to match the reference values, with the amount of adjustment defined as the calibration gain. These calibration gains were then applied to all test stimuli prior to them being processed by either coding strategy.

An overall number of ten OLSA lists and ripple discrimination sets were executed for two unilateral (InterlACE, reference ACE), and three bilateral strategy variations (InterlACE linked, InterlACE alternatively linked, reference ACE). For all experiments, five to ten-minute breaks were foreseen on the request of the subject. With one condition running approximately ten minutes, with delays already taken into account, plus two breaks of ten minutes, the test ran for an average of around two hours per experiment.

E. Analyses

Analyses included a nonparametric Wilcoxon signed-rank test because of the small sample size, the dependent relation, and impossibility of assuming a normal distribution of variables. A Benjamini and Hochberg adjustment to control the false discovery rate (FDR) within the dataset was performed [47], [48]. The linear relation between unilateral and bilateral subjective influences was assessed by the correlation of speech and spectral ripple discrimination thresholds and calculated using Pearson's correlation coefficients.

III. Results

A. Speech Intelligibility

Speech intelligibility measures with the OLSA in background noise showed no statistical significance for SRT in dB between ACE and InterlACE in the unilateral conditions and between the bilateral conditions. However, significant differences were found between unilateral and bilateral test conditions. Nevertheless, lower SRT outcomes, thus increased performances for InterlACE in both coding variations over ACE, were observed. Fig. 5 (top) represents grouped and individual speech intelligibility outcomes (SRT in dB SNR) for all assessed coding variations. With a median SRT of $-2.55\,\mathrm{dB}$, ACE unilateral showed the lowest performance. InterlACE unilateral resulted in a median SRT of $-3.25\,\mathrm{dB}$. ACE bilateral showed a median SRT of $-3.97\,\mathrm{dB}$ and bilateral InterlACE linked a median SRT of $-4.45\,\mathrm{dB}$. The largest intelligibility outcome was found in the bilateral InterlACE alternatively linked version with a median SRT of $-5.1\,\mathrm{dB}$. Both ACE and InterlACE showed improvements, with $1.42\,\mathrm{dB}$ for ACE unilateral to ACE bilateral, $1.2\,\mathrm{dB}$ for InterlACE unilateral to bilateral InterlACE linked, and $1.85\,\mathrm{dB}$ for bilateral InterlACE alternatively linked. The overall inter-quartile range (IQR) throughout all conditions showed values between $1.4\,\mathrm{dB}$ for ACE bilateral and $2.45\,\mathrm{dB}$ for InterlACE unilateral. ACE coding showed significant differences with $p \leqslant 0.05$ between its unilateral and all bilateral conditions. Similar to ACE, InterlACE unilateral resulted in significant differences compared to all bilateral conditions with $p \leqslant 0.05$.

Individual average outcomes, presented as lines in Fig. 5 (top), revealed large variabilities in performance. Two subjects CI04 and CI05 showed improvements up to $2.08\,\mathrm{dB}$ with InterlACE unilateral, while CI01, CI02, and CI03 showed comparable performance in the unilateral conditions. With the bilateral InterlACE coding variations, three subjects CI01, CI04, and CI05 improved their performance by up to $1.73\,\mathrm{dB}$, however, subjects CI02 and CI03 showed a decrease of up to $1.37\,\mathrm{dB}$.

B. Spectral Resolution

Large spectral resolution differences between grouped and individual unilateral and bilateral conditions were observed. Significant differences between ACE and InterlACE in both unilateral and bilateral conditions were observed. InterlACE showed a lower ripple discrimination performance than ACE

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2023.3322348

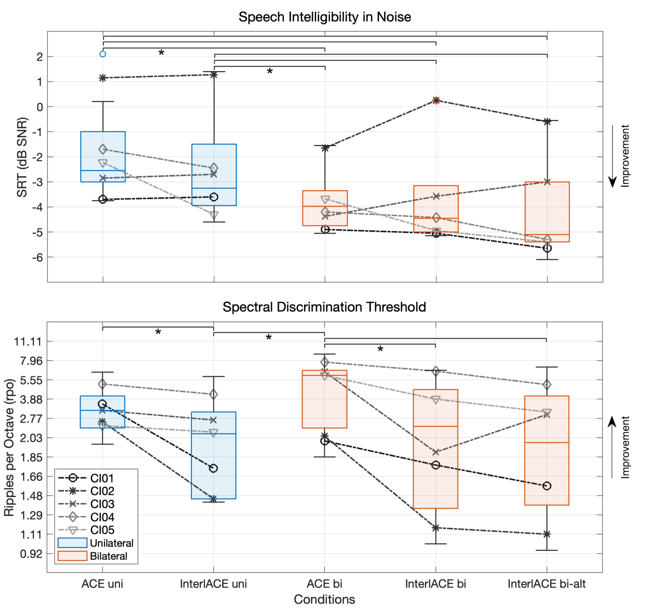

Fig. 5: Speech reception thresholds (SRT in dB SNR, top) and spectral ripple discrimination thresholds (rpo, bottom) per strategy, and average per subject.

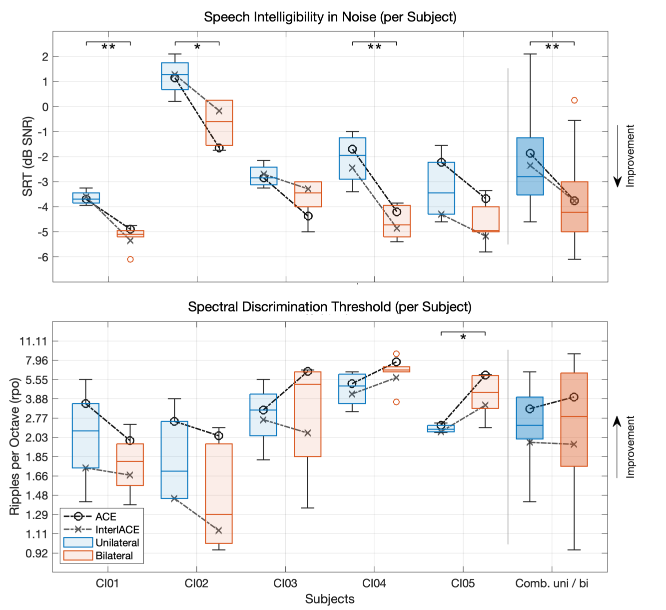

Fig. 6: Speech reception thresholds (SRT in dB SNR, top) and spectral ripple discrimination thresholds (rpo, bottom) per subject, and average per strategy, divided into unilateral and bilateral outcomes.

over all conditions. Fig. 5 (bottom) presents the ripple discrimination thresholds (rpo) for all assessed coding variations. ACE unilateral resulted in similar outputs for all five subjects with a median between 2.77 rpo and 3.88 rpo on the logarithmically spaced ripple scale. 50% of all values visible by the IQR fall in a range of around 1.5 rpo for this condition. This outcome was significantly different from the discrimination threshold observed for InterlACE and the unilateral coding condition with a $p \leqslant 0.05$ and a median threshold between 2.03 rpo and 2.77 rpo, thus a decrease in spectral resolution with the difference in median of around 1 rpo in comparison to the ACE coding unilateral. The IQR for the InterlACE unilateral condition showed a similar spread of around 1.5 rpo comparable to ACE unilateral. The ripple discrimination for the ACE bilateral condition with a median threshold between 5.55 rpo and 7.96 rpo showed a similar but more distinct discrimination behavior than observed for bilateral InterlACE with a decreased discrimination performance and median thresholds between 2.03 rpo and 2.77 rpo for the InterlACE linked condition and between 1.85 rpo and 2.03 rpo for InterlACE alternatively linked. A similar decrease was found between ACE bilateral and both InterlACE conditions with significant $p \leqslant 0.05$ outcomes. The IQR of the ACE bilateral condition showed perceived thresholds between 2.03 rpo and 7.96 rpo (range of 5.93 rpo). InterlACE in its bilateral variations showed IQRs between 1.29 rpo and 5.55 rpo (a range of 4.26 rpo).

Individual average outcomes, presented as lines in Fig. 5 (bottom), revealed a negative trend in ripple discrimination performance for all subjects in all unilateral and bilateral conditions. While the unilateral performance between ACE and InterlACE was comparable for subjects CI03, CI04, and CI05, subjects CI01 and CI02 showed the largest decrease. In the bilateral conditions CI01, CI04, and CI05 performed similarly. CI02 showed the lowest performance, while CI03 dropped in performance compared to ACE but improved between InterlACE linked and InterlACE alternatively linked coding.

C. Speech Intelligibility versus Spectral Resolution

Grouped individual unilateral and bilateral outcomes for speech intelligibility in noise and spectral discrimination thresholds are shown in Fig. 6. The results for the individual study subjects are presented similarly to the combined plot of Fig. 5 with speech-in-noise outcomes (top), and the ripple discrimination thresholds (bottom). The relation of subject performance can be compared by a visual observation of responses in both experiments.

In a comparison of grouped individual performance, Subject CI01 performed best in the speech-in-noise task with unilateral and bilateral median SRTs of -3.7 dB and -5.1 dB ($p \leqslant 0.01$) but only showed average performance for ripple discrimination. Subject CI02 showed the lowest speech performance with median SRTs of 1.27 dB and -0.6 dB with a significant improvement between unilateral and bilateral coding ($p \leqslant 0.05$), but lowest median ripple discrimination thresholds between 1.85 rpo and 1.66 rpo unilaterally and 1.29 rpo bilaterally. Subject CI04 showed the best outcomes for the ripple discrimination task with median thresholds between 5.55 rpo and 3.88 rpo unilaterally, and between 7.96 rpo and 5.55 rpo bilaterally, but an average performance in speech-in-noise compared to the other subjects, however, with significant improvement between the unilateral and bilateral condition ($p \leqslant 0.01$). Subjects CI03 and CI05 showed comparable unilateral and bilateral outcomes in speech intelligibility with median SRTs between -2.85 dB and -4.95 dB, and also performed in a similar range in the ripple discrimination task between 5.55 rpo and 2.03 rpo. Subject CI05 increased significantly in ripple discrimination ($p \leqslant 0.05$). All subjects showed differences between unilateral and bilateral conditions with an increased speech-in-noise but mixed ripple discrimination performance. On average, speech-in-noise increased

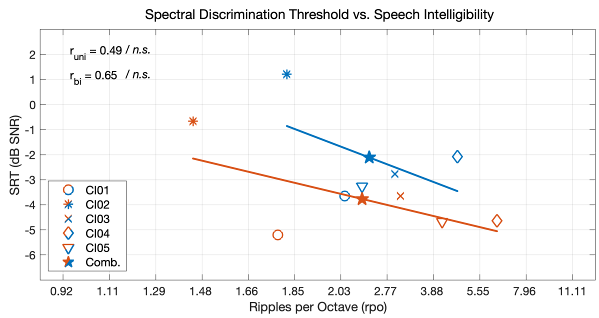

Fig. 7: Speech reception thresholds(SRT in dB SNR) versus individual spectral discrimination thresholds (rpo) for all assessed coding strategies and variations combined. Linear regressions are represented separately for unilateral (blue) and bilateral (red) outcomes.

from -2.8 dB unilaterally to -4.23 dB bilaterally ($p \leqslant 0.01$), and ripple discrimination improved from a median between 2.77 rpo and 2.03 rpo unilaterally to 3.88 rpo and 2.77 rpo bilaterally.

An individual comparison for average speech-in-noise outcomes per strategy, represented as lines in Fig. 6 (top), showed larger improvements between unilateral and bilateral conditions for ACE compared to InterlACE in subjects CI02 and CI03 between -0.94 dB and -1.35 dB, but similar unilateral to bilateral improvements for all other subjects between -1.75 dB and -2.5 dB. The same comparison for mean ripple discrimination thresholds, represented as lines in Fig. 6 (bottom), showed an overall elevated performance with similar improvements between ACE and InterlACE for all subjects. All subjects showed large differences in unilateral and bilateral outcomes with lower performance for InterlACE. CI01 and CI02 showed the largest performance gaps between ACE and InterlACE. CI04 and CI05 increased similarly for both strategies. CI03 showed a different performance with a bilateral increase with ACE but decrease with InterlACE.

To show the relation of subjective outcomes for the unilateral and bilateral combination of all InterlACE variations and reference ACE coding, a correlation of individual outcomes was performed and the Pearson's correlation coefficient r and p-values calculated, presented in Fig. 7. Intelligibility results were calculated as arithmetic means of the subject outcomes and presented as ripple discrimination thresholds (rpo) on the x-axis and SRTs (dB SNR) on the y-axis. Unilateral and bilateral correlations of ($r_{uni} = 0.49$, $p = 0.39$ and $r_{bi} = 0.65$, $p = 0.23$) resulted in a low and nonsignificant positive effect, representing a negligible, respectively, weak positive linear correlation [49]. The unilateral and bilateral regression functions showed similar slopes shifted by the overall increased individual performances for bilateral speech-in-noise.

IV. Discussion

For the assessed acute experimental speech-in-noise listening situations of this study, a favorable trend for InterlACE over ACE coding was observed, as shown in Fig. 5 (top). Two out of five subjects improved their SRTs in the unilateral conditions, three out of five subjects in the bilateral conditions. All assessed bilateral listening conditions exceeded the ACE coded outcomes similar to previous studies with bilaterally linked processors [17], [18]. The BINOM strategy was based on a link in the channel selection block of the processing with a shared clock. Outcomes showed a slight benefit compared to the bilaterally unlinked ACE processing [17]. A follow-up study compared the linkage of bilateral CI channels and processor timing in more detail and found significant differences for various conditions [18]. When the stimulation was performed synchronized between left and right processors, speech intelligibility improved over the unsynchronized method. A linkage of channels further increased the speech scores. In contrast to the speech-in-noise task with perpendicularly placed signals for assessing the performance of BINOM, the current study evaluated intelligibility performance with adaptively mixed speech-in-noise. InterlACE with its synchronized timing and linked but alternating channel selection showed a similar benefit in the speech-in-noise task, therefore supports the previously found outcomes. Even though the outcomes of the studies showed that linking the processors with the same clock improves speech-in-noise it is surprising that the interlaced selection of channels as it is performed for InterlACE coding shows only a small additional benefit in both evaluated coding variations compared to BINOM and its continuously linked variations. It seems that the method of fast switching active channels in order to gain more signal content and reduce the occurring electric crosstalk does not lead to significant benefits, at least within the speech-in-noise task and this study's small group of subjects. Hence, the results for InterlACE need to be confirmed with a larger subject group, as it was carried out for BINOM.

The ripple discrimination result showed a relatively high discrimination ability for all five study participants, as shown in Fig. 5 (bottom). Median discrimination thresholds peaked between 5.55 rpo and 7.96 rpo for the ACE bilateral coding condition. The lowest median thresholds were found for ACE bilateral and InterlACE bilateral coding and both strategy variations between 1.85 rpo and 2.77 rpo. Huge subjective differences for all assessed coding strategy variations were observed. Performance increase from unilateral to bilateral coding conditions was expected but only observed for the ACE coding. The bilateral benefit of ACE coding was not significant but a trend of around 2.5 rpo discrimination improvement between the median threshold outcomes with the peak at around 5.55 rpo was observed. Discrimination outcomes for the InterlACE coding variations showed an overall comparable spread of perceived thresholds, however, with a similar median performance between unilateral and bilateral coding variations centered at around 2 rpo.

As mentioned in several previous studies, speech intelligibility seems to directly correlate with individual spectral resolution [6], [42], [50], [51]. To review this effect for the current study, individual intelligibility and discrimination thresholds were correlated, presented in Fig. 7. No significant effects could be observed for any of the subjects and combined individual outcomes, in contrast to the previous studies. The low Pearson's correlation coefficients and nonsignificant p-values for the unilateral and bilateral conditions indicate no or only a low positive correlation between speech-in-noise and spectral resolution.

The large difference between the coding strategies and the ripple discrimination thresholds could be a result of the difference in the channel selection methodology. The ACE coding selects a defined number of maxima per processing frame, which in this study was set to eleven maxima. InterlACE selects a maximum of eleven interleaved channels per frame as well. However, the major difference is that this selection results in the activation of all available channels over two consecutive frames. ACE can potentially compensate for the missing eleven channels in the consecutive frame of stimulation for a nonstationary signal, but for the static spectrum of the ripple signal, selected signal maxima of consecutive frames will be similar over time. This effect of ACE coding will lead to a sparse variation of channels. The peaks of the signal ripples will be detected as maxima more frequently, and sites with weaker signal content will be neglected. The increased selection of maxima increases the contrast between ripples, hence, individual signal peaks can be detected easier during a ripple discrimination assessment. In contrast, the InterlACE coding does not rely on a maxima selection based on the spectrum and its intensity. Signals that elicit a channel activation will be present in the stimulation sequence, even for signal valleys, i.e. lower stimulation currents. A more uniform output sequence with similar channel activation pattern over all electrodes and reduced peak-to-valley contrast is the result. Ripple discrimination in the current experimental setup is therefore more complicated with InterlACE compared to ACE. The discussed channel selection impacts of InterlACE and ACE were investigated similar to the strategy verification of InterlACE (Fig. 4) to support this paragraph's statement.

In addition, high ripple numbers lead to aliasing effects [52]–[54]. Due to the underlying filter bank of the CI processing, the ripple resolution is limited by aliasing occurring in the output stimulation sequence with higher ripple numbers starting. Aliasing appears at lower levels in InterlACE than for ACE coding due to the increased channel activity, another possible explanation for the observed decreased ripple discrimination thresholds in the InterlACE coding variations.

No positive effects of InterlACE were found in contrast to a previous study where an increase in discrimination performance for interleaved processors was mentioned [55]. In that study, interleaving was performed between bilateral processors and twelve pitch matched pairs of electrodes, which is different to the present study. Interleaved maps were set to activate even channels on one ear and odd channels on the other ear. Hence, active right and left channels were pre-defined, in contrast to InterlACE coding of the current study, where channels are interlaced in space, time, and ear. Previous findings showed an increase in ripple discrimination thresholds for the interleaved channel activation of 1 rpo. In InterlACE coding, no pitch matching was performed. Thus, all 22 electrodes could be activated within consecutive stimulation frames and over time. This additionally could have been a factor in the decrease in discrimination performance for the unilateral and all bilateral versions of InterlACE compared to unilaterally and bilaterally linked ACE coding.

The spectral ripple discrimination assessment performed in this study showed an overall increase in perceived thresholds exceeding the results of several previous studies with thresholds between 0.62 rpo and 2.6 rpo [6], [42], [43], [56], [57]. In the current study, the median of all conditions combined resulted in thresholds around 2.77 rpo, an increase of around 1 rpo compared to reported outcomes of previous studies except for one with similar results [43], [57]. Several assumptions for the increased performance were proposed, which could as well account for the current study. The duration of implant use showed increased discrimination abilities in comparison with a previous study [6]. Another explanation was the availability of newer electrode technologies and the use of contour arrays, however, improvements have been found to be ambiguous [57], [58]. Increased ripple discrimination abilities did not only correlate with the duration of CI use but also with an increase in speech performance [57], [59], [60]. Correlations with the implant use duration or implant technologies were not explored in the current study. However, subject CI02 with the longest experience of 16.5 years showed the lowest performance in ripple discrimination and also the speech intelligibility task, contrary to the previous findings [57]. Another factor that may be important for CI02 is the subject's age of 68 years, which was highest among this group of subjects and may have affected the outcomes. This particular subject uses the MP3000 coding strategy as his clinical setup which may be an additional influential factor. Another subject, CI01, with the same electrode setup as CI02, was the youngest participant in the study with 20 years of age and a long CI experience of eight years. However, the ripple discrimination performance for CI01 was also low in comparison to CI03, CI04, and CI05. The higher performance of middle-aged subjects of this current study therefore can not be directly correlated to the duration effects found in previous studies.

The effect of number of channels and its correlation to ripple discrimination thresholds was studied before [14]. The authors split the results into two groups of good and poor performers. However, the presented function for both groups was relatively flat for both groups with a steeper increase concerning the number of active channels for the good performers. The number of active channels in the current study also showed large variations for the stimulation sequences due to the differences that occur by the channel selection paradigms between InterlACE and ACE coding. A channel count was performed by a moving average of the block length of 50 sequence frames and a block-shift of 10 sequence frames. The ripple signals showed a relatively constant activation of channels in the 0.92 rpo condition, with 10.01 electrodes active and 10.95 for the 15.74 rpo condition. For the InterlACE coded signal and the 0.92 rpo condition a mean activation of 5.24 channels was observed. The ripple signal with an increased number of ripples per octave showed similar behavior in terms of channel activations between coding strategies and signal complexity. The channel number was slightly increased for the condition of 15.74 rpo with an average of 7.19 active channels. The observed activation differences may have affected the discrimination outcomes similarly to previous findings, although the definition of the channel assignments and the resulting number of active channels was quite different in contrast to the coding performed for InterlACE and the ACE coding.

In conclusion, the hypotheses of improved speech intelligibility and spectral ripple discrimination through interlaced signal content can be at least partly supported by the present findings. With InterlACE coding, speech intelligibility in noise improved in unilateral and bilateral listening situations. These outcomes are a first indicator for the possibilities of InterlACE coding as a future clinical unilateral and bilateral speech coding strategy. A reversed effect, respectively, a performance decrease was observed for InterlACE and the spectral discrimination assessment, thus contradicting the second part of this research hypothesis. Decreased spectral resolution may have been a result of the acute testing condition. ACE coding acts as a filter to discard weak signal content, information that can be preserved in InterlACE coding. The increase in available information may have been detrimental in the first place, especially for complex signals, but may increase the discrimination performance if more training was performed. A take-home strategy version may be a solution to overcome this lack of training and would eventually increase spectral discrimination abilities. Decreased performance for the InterlACE coding in the bilateral conditions may have also been predominantly due to increased signal information. Several subjects reported noisy perceived signals with the new coding, especially in the spectral ripple discrimination task. This effect, however, may appear due to the addressed aliasing of signal information or the uniform channel selection of InterlACE with higher spectral ripple complexity. On the other hand, speech sounded very clear to the participants and did not lead to worse outcomes, shown with the speech-in-noise experiments. Overall it is worth mentioning that the small number of subjects limits a more detailed exploration of the findings of this study. However, the presented work should provide a first basis for future explorations of unilaterally and bilaterally interlaced signal content via InterlACE coding.

V. Conclusion

In this work, a novel InterlACE coding paradigm was investigated in an acute setup. By selecting even and odd-spaced channels in consecutive stimulation frames, InterlACE can activate all available electrodes. The strategy can be linked bilaterally by either activating even or even and odd electrodes synchronized in the same stimulation frame.

Speech intelligibility in background noise and spectral resolution measures were obtained and compared to ACE coding. Speech outcomes indicated improvements over ACE for two subjects unilaterally, and three subjects in the bilateral strategy variations. Furthermore, individual performance revealed larger bilateral improvements for ACE but similar improvements over all strategies.

The spectral ripple discrimination performance resulted in decreased ripple thresholds for all variations of InterlACE. The observed decrease could be due to several factors, (a) the unfamiliar acute listening situation, (b) the number of active electrodes, (c) spectral ripple aliasing effects, and (d) subject age and implant experience. Compared to previous studies, discrimination thresholds improved, which may be attributed to newer technologies or longer implantation durations. Spectral resolution needs to be re-evaluated in more detail, perhaps with alternative experimental methods.

A major influencing factor for the performance differences of InterlACE may have been the unfamiliar channel selection methodology and bilateral synchronized coding extension.

The current findings warrant a future focus on bilateral interlaced signal content. InterlACE has the potential to improve speech intelligibility in noise, which is considered a particularly challenging listening situation for CI recipients. The uniform channel coverage allows access to greater amounts of available signal content, avoiding channel clustering based on dominant signal content, potentially beneficial for noisy or complex listening situations.

Take-home trials could facilitate adaptation to the InterlACE coding paradigm. In addition, the experimental battery could be extended by assessments of channel interaction or spatial perception to explore the performance in greater detail.

Acknowledgment

The authors would like to thank the CI recipients for their participation and patience in performing the experiments. The authors would also like to thank the colleagues from research and the CI Center of the Department of Otorhinolaryngology, University Hospital Zurich.

References

[1] B. S. Wilson and M. F. Dorman, "Cochlear implants: A remarkable past and a brilliant future," *Hearing Research*, vol. 242, no. 1-2, pp. 3–21, 2008.

[2] J. Wouters *et al.*, "Sound coding in cochlear implants: From electric pulses to hearing," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 67–80, 2015.

[3] F. G. Zeng, "Challenges in improving cochlear implant performance and accessibility," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1662–1664, 2017.

[4] R. P. Carlyon and T. Goehring, "Cochlear Implant Research and Development in the Twenty-first Century: A Critical Update," pp. 481–508, oct 2021.

[5] L. K. Holden *et al.*, "Effects of stimulation rate with the nucleus 24 ACE speech coding strategy," *Ear and Hearing*, vol. 23, no. 5, pp. 463–476, 2002.

[6] B. A. Henry *et al.*, "Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners," *The Journal of the Acoustical Society of America*, vol. 118, no. 2, pp. 1111–1121, 2005.

[7] R. Y. Litovsky *et al.*, "Spatial hearing and speech intelligibility in bilateral cochlear implant users," *Ear and Hearing*, vol. 30, no. 4, pp. 419–431, 2009.

[8] A. W. Bronkhorst and R. Plomp, "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing," *The Journal of the Acoustical Society of America*, vol. 92, no. 6, pp. 3132–3139, 1992.

[9] B. A. Henry and C. W. Turner, "The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners," *The Journal of the Acoustical Society of America*, vol. 113, no. 5, pp. 2861–2873, 2003.

[10] M. F. Dorman *et al.*, "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs," *The Journal of the Acoustical Society of America*, vol. 102, no. 4, pp. 2403–2411, 1997.

[11] K. E. Fishman *et al.*, "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," *Journal of Speech, Language, and Hearing Research*, vol. 40, no. 5, pp. 1201–1215, 1997.

[12] L. M. Friesen *et al.*, "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp. 1150–1163, 2001.

[13] C. Garnham *et al.*, "Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets," *Ear and Hearing*, vol. 23, no. 6, pp. 540–552, 2002.

[14] N. B. H. Croghan *et al.*, "Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility," *The Journal of the Acoustical Society of America*, vol. 142, no. 6, pp. EL537–EL543, 2017.

[15] R. S. Tyler *et al.*, "Speech perception and localization with adults with bilateral sequential cochlear implants," *Ear and Hearing*, vol. 28, no. SUPPL.2, 2007.

[16] K. C. Hughes and K. L. Galvin, "Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing," *Cochlear Implants International*, vol. 14, no. 3, pp. 121–129, 2013.

[17] T. Gajecki and W. Nogueira, "A Synchronized Binaural N-of-M Sound Coding Strategy for Bilateral Cochlear Implant Users," in *Speech Communication; 13th ITG-Symposium*. VDE, 2018, pp. 1–5.

[18] ——, "The effect of synchronized linked band selection on speech intelligibility of bilateral cochlear implant users," *Hearing Research*, vol. 396, pp. 1–11, 2020.

[19] W. A. Ainsworth, "Perception of speech sounds with alternate formants presented to opposite ears," *Journal of the Acoustical Society of America*, vol. 63, no. 5, pp. 1528–1534, 1978.

[20] T. Lunner *et al.*, "8-channel digital filter bank for hearing aid use: Preliminary results in monaural, diotic and dichotic modes," in *Scandinavian Audiology, Supplement*, vol. 22, no. 38, 1993, pp. 75–81.

[21] D. S. Chaudhari and P. C. Pandey, "Dichotic presentation of speech signal with critical band filtering for improving speech perception," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 6, 1998, pp. 3601–3604.

[22] A. N. Cheeran and P. C. Pandey, "Evaluation of Speech Processing Schemes using Binaural Dichotic Presentation to Reduce the Effect of Masking in Hearing-Impaired Listeners," in *18th International Congress on Acoustics*, Kyoto, Japan, 2004, pp. 1523–1526.

[23] P. N. Kulkarni *et al.*, "Binaural dichotic presentation to reduce the effects of spectral masking in moderate bilateral sensorineural hearing loss," *International Journal of Audiology*, vol. 51, no. 4, pp. 334–344, 2012.

[24] D. Lawson *et al.*, "Fourth Quarterly Progress Report. Speech processors for auditory prostheses," NIH project N01-DC-8-2105, Tech. Rep., 1999.

[25] R. S. Tyler *et al.*, "An attempt to improve bilateral cochlear implants by increasing the distance between electrodes and providing complementary information to the two ears," *Journal of the American Academy of Audiology*, vol. 21, no. 1, pp. 52–65, 2010.

[26] P. N. Kulkarni *et al.*, "Perceptually balanced filter response for binaural dichotic presentation to reduce the effect of spectral masking," *The Journal of the Acoustical Society of America*, vol. 120, no. 5, pp. 3253–3253, 2006.

[27] R. V. Shannon, "A Model of Temporal Integration and Forward Masking for Electrical Stimulation of the Auditory Nerve," *Cochlear Implants*, pp. 187–205, 1990.

[28] B. Swanson and H. Mauch, "Nucleus MATLAB Toolbox 4.20 Software User Manual," Cochlear Ltd., Lane Cove NSW, Australia, Tech. Rep., 2006.

[29] J. Poppeliers, "NIC3 V1.2 Streamer Release Report," Cochlear Ltd., Lane Cove NSW, Australia, Tech. Rep., 2014.

[30] F. G. Zeng *et al.*, "Cochlear Implants: System Design, Integration, and Evaluation," *IEEE Reviews in Biomedical Engineering*, vol. 1, pp. 115–142, 2008.

[31] C. A. Verschuur, "Effect of stimulation rate on speech perception in adult users of the Med-El CIS speech processing strategy," *International journal of audiology*, vol. 44, no. 1, pp. 58–63, jan 2005.

[32] R. V. Shannon *et al.*, "Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise," *Audiology and Neurotology*, vol. 16, no. 2, pp. 113–123, 2011.

[33] M. J. Shader *et al.*, "The effect of stimulation rate on speech understanding in older cochlear-implant users," *Ear and hearing*, vol. 41, no. 3, p. 640, may 2020.

[34] A. E. Vandali *et al.*, "Speech perception as a function of electrical stimulation rate: Using the nucleus 24 cochlear implant system," *Ear and Hearing*, vol. 21, no. 6, pp. 608–624, 2000.

[35] K. Plant *et al.*, "Clinical evaluation of higher stimulation rates in the nucleus research platform 8 system," *Ear and Hearing*, vol. 28, no. 3, pp. 381–393, 2007.

[36] B. Hagerman, "Sentences for Testing Speech Intelligibility in Noise," *Scandinavian Audiology*, vol. 11, no. 2, pp. 79–87, 1982.

[37] K. Wagener *et al.*, "Entwicklung und Evaluation eines Satztests in deutscher Sprache. Teile I, II, und III (in German). (Development and evaluation of a German sentence test–Part I, II, and III)," *Zeitschrift für Audiologie*, vol. 38, pp. 4–15, 44–56, 86–95, 1999.

[38] ——, "Design, optimization and evaluation of a Danish sentence test in noise," *International Journal of Audiology*, vol. 42, no. 1, pp. 10–17, 2003.

[39] K. Wagener, *Factors influencing sentence intelligibility in noise*. BIS Verlag, 2004.

[40] B. Kollmeier *et al.*, "The multilingual matrix test: Principles, applications, and comparison across languages: A review," *International Journal of Audiology*, vol. 54, pp. 3–16, 2015.

[41] T. Brand and B. Kollmeier, "Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests," *The Journal of the Acoustical Society of America*, vol. 111, no. 6, pp. 2801–2810, 2002.

[42] J. H. Won *et al.*, "Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users," *JARO - Journal of the Association for Research in Otolaryngology*, vol. 8, no. 3, pp. 384–392, 2007.

[43] W. R. Drennan *et al.*, "Validation of a clinical assessment of spectral-ripple resolution for cochlear implant users," *Ear and Hearing*, vol. 35, no. 3, 2014.

[44] H. Levitt, "Transformed up-down methods in psychoacoustics," *Journal of the Acoustical Society of America*, vol. 49, no. 2, pp. Suppl 2:467+, 1971.

[45] S. Tabibi, "A bio-inspired coding (BIC) strategy for cochlear implants - PhD Thesis," Ph.D. dissertation, ETH, 2020.

[46] W. K. Lai and N. Dillier, "RFcap: A software analysis tool for multichannel cochlear implant signals," *Cochlear Implants International*, vol. 14, no. 2, pp. 107–116, 2013.

[47] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

[48] Y. Benjamini and Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

[49] J. Cohen, "A power primer," *Psychological Bulletin*, vol. 112, no. 1, pp. 155–159, 1992.

[50] L. K. Holden *et al.*, "Factors affecting outcomes in cochlear implant recipients implanted with a perimodiolar electrode array located in scala tympani," *Otology and Neurotology*, vol. 37, no. 10, pp. 1662–1668, nov 2016.

[51] M. Lawler *et al.*, "Comparison of the spectral-temporally modulated ripple test with the Arizona biomedical institute sentence test in cochlear implant users," *Ear and Hearing*, vol. 38, no. 6, pp. 760–766, 2017.

[52] M. Winn and G. O'Brien, "Flaws in Spectral Ripple Stimuli for Listeners with Cochlear Implants," *PsyArXiv*, 2019.

[53] J. M. Resnick *et al.*, "Spectral aliasing in an acoustic spectral ripple discrimination task," *The Journal of the Acoustical Society of America*, vol. 147, no. 2, pp. 1054–1058, feb 2020.

[54] M. B. Winn and G. O'Brien, "Distortion of Spectral Ripples through Cochlear Implants Has Major Implications for Interpreting Performance Scores," *Ear and Hearing*, vol. 43, no. 3, pp. 764–772, may 2022.

[55] J. M. Aronoff *et al.*, "Interleaved Processors Improve Cochlear Implant Patients' Spectral Resolution," *Ear and Hearing*, vol. 37, no. 2, pp. e85–e90, 2016.

[56] E. S. Anderson *et al.*, "Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users," *The Journal of the Acoustical Society of America*, vol. 130, no. 1, pp. 364–375, 2011.

[57] E. K. Jeon *et al.*, "Cochlear implant users' spectral ripple resolution," *The Journal of the Acoustical Society of America*, vol. 138, no. 4, pp. 2350–2358, 2015.

[58] L. T. Cohen *et al.*, "Psychophysical measures in patients fitted with Contour™ and straight Nucleus electrode arrays," *Hearing Research*, vol. 212, no. 1-2, pp. 160–175, 2006.

[59] J. Z. Sarant *et al.*, "Variation in speech perception scores among children with cochlear implants," *Ear and Hearing*, vol. 22, no. 1, pp. 18–28, 2001.

[60] L. S. Davidson *et al.*, "Factors contributing to speech perception scores in long-term pediatric cochlear implant users." *Ear and hearing*, vol. 32, no. 1 Suppl, 2011.