



HHS Public Access

Author manuscript

IEEE/ACM Trans Audio Speech Lang Process. Author manuscript; available in PMC 2022 January 01.

Published in final edited form as:

IEEE/ACM Trans Audio Speech Lang Process. 2021 ; 29: 265–273. doi:10.1109/taslp.2020.3039601.

The Temporal Limits Encoder as a Sound Coding Strategy for Bilateral Cochlear Implants

Alan Kan [Member, IEEE],

Waisman Center, University of Wisconsin-Madison at the time this work was conducted. He is now with the School of Engineering, Macquarie University, NSW, Australia, 2109

Qinglin Meng [Member, IEEE]

Acoustics Laboratory, School of Physics and Optoelectronics, South China University of Technology, Guangzhou, China, 510641

Abstract

The difference in binaural benefit between bilateral cochlear implant (CI) users and normal hearing (NH) listeners has typically been attributed to CI sound coding strategies not encoding the acoustic fine structure (FS) interaural time differences (ITD). The Temporal Limits Encoder (TLE) strategy is proposed as a potential way of improving binaural hearing benefits for CI users in noisy situations. TLE works by downward-transposition of mid-frequency band-limited channel information and can theoretically provide FS-ITD cues. In this work, the effect of choice of lower limit of the modulator in TLE was examined by measuring performance on a word recognition task and computing the magnitude of binaural benefit in bilateral CI users. Performance listening with the TLE strategy was compared with the commonly used Advanced Combinational Encoder (ACE) CI sound coding strategy. Results showed that setting the lower limit to 200 Hz maintained word recognition performance comparable to that of ACE. While most CI listeners exhibited a large binaural benefit (6 dB) in at least one of the conditions tested, there was no systematic relationship between the lower limit of the modulator and performance. These results indicate that the TLE strategy has potential to improve binaural hearing abilities in CI users but further work is needed to understand how binaural benefit can be maximized.

Keywords

assistive technology; auditory implant; bilateral cochlear implant; signal processing; sound coding strategy; speech processing

I. Introduction

THE cochlear implant (CI) has been remarkably successful in restoring speech understanding to the profoundly deaf. This is achieved by converting acoustic sound into pulsatile electrical stimulation that stimulates the auditory nerve. While there are some differences in signal processing steps between CI manufacturers (see [1] for more details),

modern CIs use Continuous Interleaved Sampling [2] as a basis for signal conversion. The conversion process begins by separating the incoming sound into a small number of frequency bands (12 to 22, depending on manufacturer). Then, in each band, the low-passed filtered signal envelope (ENV) is extracted and used to amplitude-modulate a high-rate train of electrical pulses (typically 900 pulses per second per channel). The temporal fine structure (TFS) of the signal in each channel is typically not encoded. The amplitude-modulated pulses are then sent to an implanted electrode array to stimulate different parts of the cochlea; taking advantage of the place-to-frequency mapping of the cochlea to transmit the different frequencies of a sound to the brain. To reduce current interaction among electrode channels, the pulses are generated without temporal overlap at different electrodes, i.e., in an interleaved manner. Using this form of processing, patients who are deaf have been able to recover a relatively large amount of speech understanding; especially in quiet situations, where some CI users can achieve 100% speech recognition without lip-reading [3]. However, in noisy situations, performance is still much poorer than that of normal-hearing (NH) listeners [4], [5].

Bilateral implantation (i.e., providing a CI in each ear) has been shown to provide significant benefits for speech understanding in noise when compared to using only one CI [6], [7]. With bilateral implantation, the auditory system can take advantage of spatial hearing cues to provide a binaural hearing benefit. For a sound located on the horizontal plane, there are two important spatial hearing cues: (1) the interaural time difference (ITD) which arises due to a path length difference of arrival of a sound to each ear; and (2) the interaural level difference (ILD) which arises due to the head acting as an acoustic barrier causing a difference in the sound level at the two ears. ITDs are the dominant cue for sound localization at low frequencies (< 1500 Hz) and are decoded mainly from the TFS of the acoustic signal by the auditory system, while ILDs are the dominant cue frequencies ≥ 4000 Hz [8]. These two cues are used by the brain to locate the direction of a sound in space and to improve speech understanding in noisy situations.

Despite improved outcomes with bilateral implantation, there is still a gap in performance for speech understanding in noise when compared with NH listeners [9]. This gap in performance has been attributed to a lack of reliance on ITDs for locating the direction of a sound in bilateral CI users [10], [11]. While there are a number of clinical and individual reasons why bilateral CI users are unable to use ITDs (see [12] for a detailed review), one major factor is that the signal processing does not encode all the acoustic cues necessary for the brain to take advantage of having two input signals. In particular, ITDs in the TFS of the acoustic signal are typically not encoded by CIs. Attempts to improve the effectiveness of bilateral CIs have focused on developing bilateral sound coding strategies that will provide the acoustic TFS information by encoding it in the timing of the electrical pulses [13]–[16]. In these strategies, the firing of electrical pulses on electrodes associated with low frequency content are timed to certain features of the acoustic TFS, such as the signal peak or zero crossing. These strategies also reduce the firing rate in these low frequency channels. There are two reasons for reducing the stimulation rate on these channels: (1) ITDs are a low frequency cue [8]; and (2) psychophysical studies with bilateral CI users have shown that sensitivity to ITDs is better at low (around 100 pulses per second), rather than high, rates of stimulation [17]. However, lowering the stimulation rate can potentially reduce speech

recognition performance due to envelope under-sampling [15]. To date, these mixed-rate strategies have only led to small improvements in speech understanding in noise [18], [19]; suggesting that alternative approaches should be considered to improve outcomes with bilateral CIs.

II. Temporal Limits Encoder

The Temporal Limits Encoder (TLE) strategy was originally proposed to enhance the representation of TFS cues for improving pitch discrimination [20], Mandarin tone recognition, and speech-in-noise recognition with unilateral CIs [21]. The TLE strategy works by down-transposition of mid-frequency band-limited channels to a lower intermediate frequency that is within the range of good temporal sensitivity for pitch (Fig. 1a). In this manner, the TFS of each band is implicitly converted to a much slower version without explicit ENV/TFS decomposition and at a frequency range that is possible for the auditory nerves to encode [21]. When the TLE strategy is applied in both ears with synchronized pulse timing control, the ITD in the TFS of the signal is theoretically also encoded [22].

Specifically, band-limited information in the k -th channel can be expressed as a function of quasi-sinusoidal oscillations

$$\begin{aligned} s_k(t) &= e_k(t)\cos(\phi_k(t)) \\ &= e_k(t)\cos(2\pi f_{ck}t + \psi_k(t)) \end{aligned} \quad (1)$$

where, $e_k(t)$ and $\cos(\phi_k(t))$ terms denote the ENV and TFS of the channel, respectively. Here, f_{ck} can be any frequency value because $\psi_k(t)$ is a function that can compensate for different selection of f_{ck} .

Frequency down-transposition of $s_k(t)$ can be achieved by $s_k(t) \times \cos(2\pi f_m t)$ followed by a low-pass filter. If the transposition range $f_m = f_{ck} - (f_{low} + f_{up})/2$, then this process yields a frequency down-transposed replica of $s_k(t)$

$$u_k(t) = e_k(t)\cos(\psi_k(t)) \quad (2)$$

where, f_{low} and f_{up} denote the lower and upper frequency bounds of $s_k(t)$ respectively. Here, $u_k(t)$ comprises of the original ENV and a slower version of the TFS of the channel. $u_k(t)$ can then be used as the amplitude modulator of the k -th channel's electric pulses in TLE. It should be noted that $s_k(t)$ and $u_k(t)$ have the same Hilbert envelope (used as the amplitude modulator in CIS) and the same spectral structure. The difference between $s_k(t)$ and $u_k(t)$ is their spectral centroids. They can be perfectly converted to each other as long as f_m is chosen appropriately. In TLE, f_m is chosen to be

$$f_m = f_{low} - f_{lim}, \quad (3)$$

where f_{lim} denotes a user-defined parameter of the lower limit of temporal pitch perception at single electrode. Consequently, f_{lim} is also the lower frequency bound of $u_k(t)$. It is suggested that f_{lim} be higher than 50 Hz [21]; otherwise, the modulator may contain

components at very low frequencies (<50 Hz) where they are rarely perceived as pitch variations by CI users [23]. Hence, for TLE, an intermediate frequency should be chosen for f_m . With this processing, the overall signal envelope in the band-limited channel is maintained but places the TFS at a range that is within the limits of pitch perception and ITD sensitivity.

In the bilateral CI case (Fig. 1b), where the sampling and electrical pulse generation clocks of the processors in each ear are assumed to be synchronized, the signals arriving at the two implant microphones are

$$\begin{aligned} s_k^L(t - \tau) &= e_k^L(t - \tau) \cos(\phi_k^L(t - \tau)) \\ &= e_k^L(t - \tau) \cos(2\pi f_{ck}(t - \tau) + \psi_k^L(t - \tau)) \\ &= e_k^L(t - \tau) \cos(2\pi f_{ck}(t) + \psi_k^L(t - \tau) - 2\pi f_{ck}\tau) \\ s_k^R(t) &= e_k^R(t) \cos(\phi_k^R(t)) \\ &= e_k^R(t) \cos(2\pi f_{ck}(t) + \psi_k^R(t)) \end{aligned} \quad (4)$$

where τ denotes the ITD. The difference in band-limited information across the ears will be down-modulated by TLE to yield

$$\begin{aligned} u_k^L(t - \tau) &= e_k^L(t - \tau) \cos(\psi_k^L(t - \tau) - 2\pi f_{ck}\tau) \\ u_k^R(t) &= e_k^R(t) \cos(\psi_k^R(t)) \end{aligned} \quad (5)$$

In terms of binaural cues, ILDs are preserved by the TLE transposition because the overall signal envelopes are left unchanged. That is,

$$\text{ILD} = 20 \log_{10} \frac{e_k^R(t)}{e_k^L(t - \text{ITD})} \quad (6)$$

in decibels. Further, according to (4) and (5), the interaural phase difference (IPD) is also unchanged through the TLE transposition:

$$\text{IPD} = \psi_k^R(t) - \psi_k^L(t - \text{ITD}) + 2\pi f_{ck} \text{ITD}, \quad (7)$$

which implies IPD cues in the signal are preserved in the TFS of TLE modulators. In terms of ITDs, for a pure tone sound source with frequency f , the original ITD between the TFS of $s_k^R(t)$ and $s_k^L(t - \tau)$ is

$$\text{ITD}_s = \frac{\text{IPD}}{2\pi f}, \quad (8)$$

while for the actual ITD between the TFS of the TLE modulators (i.e. $u_k^R(t)$ and $u_k^L(t - \tau)$) is

$$\text{ITD}_u = \frac{\text{IPD}}{2\pi(f - f_m)} = \text{ITD}_s \frac{f}{f - f_m}. \quad (9)$$

Hence, ITD_s is constant across frequency channels while ITD_u is dependent on f_m or f_{ck} (which is different across channels as defined by (3)) and its relative difference to the signal frequency f . Therefore, while the IPD is preserved between TLE modulators at two ears, the frequency transposition of TLE may lead to different ITDs across different electrodes. Even so, TLE is expected to provide useful binaural information because of the preserved IPD.

Compared with previously proposed bilateral sound coding strategies, the TLE strategy is novel in that it does not rely on an explicit extraction and encoding of a “feature” of the TFS such as a peak or zero-crossing. Hence, implementation is simplified. Further, the rate of stimulation is not explicitly lowered, which should help maintain speech understanding performance comparable to that of modern CI sound coding strategies.

Listening tests with the TLE strategy have shown some promising results for enhancing Mandarin speech understanding and improving binaural hearing benefits. In NH listeners using CI vocoder simulations, improvements have been shown for both Mandarin tone discrimination and speech recognition in noise with the TLE strategy [21], [22]. In particular, the TLE strategy was able to provide an average of 4 dB binaural hearing benefit for understanding speech in noise to NH listeners, though a higher signal to noise ratio (SNR) was needed in the reference condition when compared to listening with the continuous interleaved sampling strategy [22]. In bilateral CI users, listening tests have shown that word recognition in quiet with the TLE strategy was comparable to that of the Advanced Combinational Encoder (ACE) strategy (used in Cochlear Ltd Nucleus sound processors) with only a short period of acclimatization [24]. Average word recognition scores for TLE and ACE strategies were 86.5% vs 90.3%, respectively. Subjective reports from CI users listening with the TLE strategy indicated an initial noticeable difference in voice quality (lower pitch) compared to ACE. However, this gradually sounded more “normal” after conversing for an hour while listening with the TLE strategy. For word recognition in noise, the TLE strategy was able to provide a modest 1.3 dB additional binaural hearing benefit when compared to listening with the ACE strategy in the bilateral CI users tested.

While the TLE strategy is able to provide improved binaural hearing benefits to bilateral CI users, the amount of benefit varied greatly between the CI users. In [24], five out of eight listeners had a total binaural hearing benefit ranging from 3 to 11 dB. There were two listeners who showed no benefit with the TLE strategy and one had a decrement in performance. One possible reason for the across-listener variance is that the parameter f_{lim} was fixed at 150 Hz for the experiment in [24], which may not have been an optimal choice. In this work, we investigated the effect of choice of f_{lim} on word recognition in bilateral CI users.

III. Experiment

A. Listeners

Seven, postlingually-deafened bilateral CI users (6 females, 1 male) participated in the experiments. All were native speakers of American English and used the ACE strategy in everyday listening. Listeners were between the ages of 54 and 78 with an average age of 69

years, and had at least 7 years of bilateral CI experience. All listeners had participated in the TLE strategy experiment described in [24], and were sensitive to ITDs presented via direct electrical stimulation as tested using methods described in [11]. Testing was conducted at the University of Wisconsin-Madison and listeners were paid a daily stipend for their time along with reimbursements for travel costs to Madison, Wisconsin. Experimental procedures conformed to the regulations set by the National Institutes of Health and were approved by the University of Wisconsin-Madison's Health Science Institutional Review Board.

B. Setup

The experiments were conducted in a single-walled sound booth which had additional sound absorbing foam attached to the inside walls to reduce reflections. The booth houses a semi-circular array of loudspeakers (Cambridge SoundWorks) with a radius of 1.2 m. Loudspeakers were connected to a Tucker-Davis Technologies System3 (with RP2.1, SA1 and PM2 modules) outside the booth and controlled via custom-written MATLAB software. For these experiments, only the loudspeakers at 0° (front), -90° (left) and $+90^\circ$ (right) were used. The listener sat in the middle of the array with their head positioned at approximately the same height as the loudspeakers. A touchscreen was placed in front of the listener at knee height for collecting responses.

C. Sound Coding Strategy Implementation

Listening tests were conducted using the CCI-MOBILE. Developed at University of Texas at Dallas, the CCI-MOBILE is a bilaterally-synchronized research platform which allows real-time testing of new sound coding strategies with CI users implanted with Cochlear Ltd CI24 family of electrode arrays. Further details about the hardware design of the CCI-MOBILE can be found in [25].

An example MATLAB implementation of the ACE strategy is provided with the CCI-MOBILE, and has been shown to give speech recognition results comparable to that of commercial Cochlear sound processors [26]. The processing steps of the ACE strategy, as implemented on the CCI-MOBILE, are shown in Figure 2a and described as follows. The same processing steps are applied to both ears.

1. The CCI-MOBILE samples incoming sound at a rate of 16 kHz and buffers the audio for 8 ms (128 samples) before passing to MATLAB.
2. Overlap-add processing is applied in 128-sample frames, and frames are shifted according to the stimulation rate (typically 900 pulses per second).
3. A 128-point Fast Fourier Transform (FFT) is applied to each frame and only the positive frequency bins are retained.
4. The remaining FFT bins are then combined into 22 channels according to the standard frequency allocation of Cochlear sound processors (see [27] for channel frequency allocations).
5. The N (typically 8) highest maxima are chosen from the 22 channels.

6. The N maxima are compressed using the function $\log(1 + bx)/\log(1 + b)$, where x is the input and b is calculated from clinical and sound processor parameters. For the CCI-MOBILE, $b = 415.995$.
7. The compressed maxima are scaled by the electric dynamic range of a listener and added to an output buffer.
8. Once 8 ms worth of pulses have been processed, the output buffer is passed to the CCI-MOBILE for electrode stimulation.

For these experiments, the example MATLAB code was modified so that the ACE and TLE strategies could be implemented within a common framework for ease of switching between strategies during testing, as well as to increase the speed of the code and support the range of clinical stimulation rates that could be used (900 to 1800 pulses per second) without pulses being dropped across 8 ms frames.

The TLE strategy, as shown in Figure 2b, was implemented on the CCI-MOBILE by adding a frequency transposition step prior to Step 4 (above) in the channels that are assigned to frequencies below ~ 1500 Hz. For a standard 22 channel mapping, TLE is applied to FFT bins 3 to 11 which maps to channels 22 to 13. The TLE strategy can be implemented in the FFT domain by multiplying each bin by $e^{j(\pi(n-1) - \theta_{k,g})}$, where n is the index of the FFT frequency bin, k is the channel number, g is the frame number, and

$$\theta_{k,g} = \theta_{k,g-1} + 2\pi f_m T_{shift}, \quad (10)$$

where T_{shift} is the time shift between adjacent frames and f_m is the transposition range as defined Eq. 3. Then, for each band, the real part is half-wave rectified to get the TLE modulator [28]. Because only the positive frequency information of the FFT is used, there is no aliasing in the TLE frequency down-transposition in the FFT domain.

Listeners were tested on four sound coding strategy conditions: the TLE strategy with f_{lim} in Eq. (3) set to 100, 200 or 300 Hz (referred to as TLE100, TLE200, and TLE300 in this paper), and the ACE strategy as implemented on the CCI-MOBILE. Figure 3 shows example electrode outputs of the ACE strategy and the three TLE conditions. It can be seen that as f_{lim} decreases, the density of pulses on low frequency channels is decreased. Further, recall that the choice of f_{lim} does not change the IPD at each electrode (see Section II) but the actual ITD introduced at each electrode will be relative to the frequency range of the TLE modulators. When compared to ACE, it can be seen that the TLE conditions introduce additional IPD and ITD cues in the TFS of the pulse trains at the low frequency channels. It should also be noted that the TLE strategy causes different electrodes to be activated compared to ACE because the change in envelope modulation at the lower frequency channels leads to higher frequency channels being selected during the maxima picking process (step 5 above). This effect can be most easily seen in electrode 4 in Fig. 3. The impact of these cues were examined using a listening test with bilateral CI users.

D. Listening Test Conditions

For the listening test, three conditions were tested: (1) Quiet condition - target talker presented from the front loudspeaker; (2) Co-located condition - target talker and two interferers, all presented from the front loudspeaker; and (3) Separated condition - target talker from the front and an interferer was presented from each of the left and right loudspeakers. The target talker was a male voice speaking a mono-syllable word from the Consonant-Nucleus-Consonant (CNC) word test [29] and the interferer was a female voice speaking sentences [30]. Word recognition in quiet was always tested first, followed by the conditions with interferers.

In the quiet condition, words were presented at a fixed level of 60 dB SPL (A-weighted). Listeners were tested on 50 words per sound coding strategy condition. Testing of the 50 words were divided into two blocks of trials, each consisting of 25 words. Blocks of trials for each sound coding strategy condition were tested in a randomly interleaved manner for each listener. In each trial, a single word was played and the listener indicated their response by choosing the word they heard from a choice of 50 words shown on the touchscreen in front of them. The choice of words shown on the touchscreen was different for each block of trials.

In the Co-located and Separated conditions, the interferer was presented at a fixed level of 50 dB SPL (A-weighted) and always started/ended 250 ms before/after the target word. The presentation level of the target word was adaptively changed on each trial using a two-down, one-up staircase with twelve turnarounds. This type of adaptive tracking is used to estimate the SNR needed to obtain 71% correct word recognition [31]. The staircase started at an initial SNR of +5 dB and moved in 3 dB steps until the fourth turnaround, where the step size was then changed to 2 dB. As in the quiet condition, the listener's task was to identify the word presented from the front loudspeaker. The order of testing of sound coding strategy and spatial (Co-located or Separated) condition combinations was randomized for each listener, and two adaptive tracks were collected for each combination. Once all adaptive tracks were collected, the SNR needed to obtain 71%-correct word recognition, also known as the speech reception threshold (SRT), was calculated for each condition from a Bayesian estimation of the psychometric function derived from the data [32].

Two conditions with interferers were included because it allows for the calculation of a binaural hearing benefit due to the spatial separation of the target talker from interferers [9]. This benefit, also known as spatial release from masking (SRM), is calculated by subtracting the performance in the Separated condition from that of the Co-located condition. This metric was used to determine the amount of binaural hearing benefit provided by the TLE strategy.

IV. Results

Percent correct word recognition in quiet for the different sound coding strategy conditions is shown in Fig. 4. Individual scores are provided in Table I. On average, listeners had highest percent correct word recognition when listening with the ACE strategy on the CCi-MOBILE. Across the three TLE conditions, mean performance declined with decreasing

f_{lim} . A repeated-measures analysis of variance (ANOVA) revealed statistically-significant differences in percent correct scores ($F_{3,18} = 3.995$, $p = 0.02$). Pairwise comparison with Bonferroni correction revealed that the scores for TLE100 was significantly different to those obtained with the ACE strategy ($p = 0.01$).

SRTs (SNR needed to obtain 71% correct word recognition with interferers) obtained in the Co-located and Separated conditions are shown in Fig. 5. Individual SRTs are provided in Table II. On average, SRTs for the Co-located condition was higher than the Separated condition for all strategies tested. SRTs were similar across the sound coding strategy conditions in each spatial configuration of target/interferer tested. Two-way, repeated-measures ANOVA revealed a significant effect of configuration of target/interferer ($F_{1,6} = 58.50$, $p < 0.001$) but not of strategy ($F_{3,18} = 0.39$, $p = 0.99$). There was no significant interaction between the two factors.

SRM, a measure of binaural benefit for speech understanding in noise, is shown in Fig. 6. On average, there was no significant difference in binaural benefit between the different sound coding strategy conditions tested ($F_{3,18} = 0.224$, $p = 0.88$). This is due to the fact that there appears to be no systematic trend in binaural benefit as a function of choice of f_{lim} in the TLE strategy at an individual level (Fig. 7). However, for each listener, the TLE strategy does appear to provide a substantial binaural benefit in at least one of the f_{lim} conditions tested. This is summarized in Table III, where the TLE strategy provides an average of 7.4 dB binaural hearing benefit. However, word recognition in quiet with these f_{lim} can be a few percentage points poorer than that of ACE.

V. Discussion

In this work, the TLE strategy's ability to improve binaural speech understanding in noise for bilateral CI users was evaluated and compared to that obtained via listening with the ACE strategy. The TLE strategy works by transposing the mid-frequency TFS to a lower frequency range, and then modulates the pulse firing of the electrodes in the original mid-frequency region to provide additional TFS information in these channels. This is different to hearing aid transposition methods which down-transposes the unhearable high frequency components to hearable middle frequency sounds by stimulating the mid-frequency cochlea region [33]. The TLE strategy does not change the cochlea region being stimulated.

For word recognition in quiet, results showed that setting f_{lim} to 200 Hz and above yielded comparable performance to that of the ACE strategy. However, setting f_{lim} to 100 Hz led to a small (~5%) but significant reduction in word recognition in quiet compared to ACE. This finding is compatible with results previously reported with the TLE strategy in [24]. In [24], f_{lim} was set to 150 Hz and word recognition scores were not significantly different to that of the ACE strategy. Taken together, these results suggest that TLE processing has minimal effect on word recognition performance if f_{lim} is set to 150 Hz. The reduction in word recognition with lower f_{lim} is somewhat to be expected given that the encoding of the speech envelope may have been compromised by the lower density of pulses in the channels where TLE was applied (see Fig. 3). Poorer speech recognition with low stimulation rates has been shown in [15], [34]. However, the drop in word recognition performance is much smaller

with TLE than that reported in prior studies. For example, [15] reported an ~20% drop in performance when the stimulation rate was lowered from around 1000 to 100 pulses per second. The difference with TLE may be that there are still more pulses encoding the channel envelope compared to that used in [15]. It should be noted that listeners in this experiment were not given an opportunity to practice listening with the different f_{lim} conditions prior to testing and it is possible that given a period of acclimatization, word recognition with lower f_{lim} settings may approach that of the ACE strategy.

When listening in noise, there was no systematic relationship between choice of f_{lim} and binaural hearing benefit. The inability to find a systematic relationship may be due to the heterogeneity of the CI user population. There are several factors that contribute to the variability in outcomes observed in bilateral CI users. For example, CI users may experience prolonged duration of deafness prior to implantation which can lead to loss of spiral ganglion cells [35]–[37]. This leads to “dead regions” along the cochlea that are insensitive to electrical stimulation and can affect the transmission of information along the electrode-neural interface [38]. For bilateral CI users, the dead regions may be different in each ear which may lead to large asymmetries in speech recognition performance between the two ears and, in turn, affect binaural hearing outcomes [39]–[41]. Further, surgical insertion depth of the electrode arrays typically vary [42], [43]. If electrode insertion depths are different in each ear, acoustic information at the same frequency may be used to stimulate different cochlea locations in the two ears leading to an interaural place-of-stimulation mismatch (IPM). IPM has been shown to affect sensitivity to ITDs and ILDs [44]–[46], and binaural benefits for speech understanding in noise [47]–[49].

It is interesting to note that most individual listeners did obtain a binaural benefit of at least 6 dB or greater in at least one of the f_{lim} conditions tested. To understand the significance of the binaural benefit with TLE, NH listeners were tested in the Co-located and Separated conditions using the same setup. Average SRTs in NH listeners were –8.5 dB and –16.2 dB for Co-located and Separated conditions, respectively, yielding an average SRM of 7.7 dB. Hence, baseline speech understanding in noise is still much poorer with bilateral CIs when listening with the TLE strategy when compared to NH listeners. Positive SNRs were needed by CI users to reach 71% correct word recognition in noise (see Table II) while NH listeners were able to achieve this level of performance at much more adverse SNRs. However, it seems that the TLE strategy has the potential to provide a binaural hearing benefit but the underlying mechanism enabling this benefit remains unclear. The present data would suggest that this benefit is not dependent on the f_{lim} parameter alone. Further work is needed to understand how individual factors might interact with f_{lim} to yield or inhibit a binaural benefit with the TLE strategy. It should also be noted that the conditions tested are not necessarily representative of real-life listening situations and that the actual benefit might be smaller than that observed in this experiment.

Compared to previously proposed bilateral sound coding strategies, the TLE strategy shows much promise. In terms of implementation, the TLE strategy is much simpler compared to those previously proposed as it does not rely on detection of acoustic features such as a peak [13] or zero-crossing [14], [15]. Rather, the TLE strategy modulates the pulse firing of mid-frequency electrodes with a down-modulated version of the mid-frequency TFS. In our

implementation, real-time processing could be achieved on the CCI-MOBILE research platform. In terms of performance, the TLE strategy is able to maintain speech understanding performance for regular users of the ACE strategy while providing at least 2 dB improvement in binaural hearing benefit. In some f_{lim} conditions, a much larger binaural benefit was observed (Table III) and substantially greater than that reported using peak-picking methods (e.g. [18], [19]), where binaural benefits were on the order of 1–2 dB only. However, it is unclear how the TLE strategy is providing this large binaural benefit. One possibility is that the binaural benefit with the TLE strategy could have arisen through the encoding of important across-frequency TFS modulation relationships that would otherwise have been lost in a feature encoding strategy. Previous work has shown the importance of TFS modulations for comodulation masking release [50], speech understanding in noise at a unilateral level [51] and melody recognition [52]. Another possibility is that ITD sensitivity has been enhanced via the TLE strategy. While psychophysical experiments have shown that low stimulation rates are important for enabling ITD sensitivity [17], the TLE strategy does not explicitly lower the stimulation rate to enhance ITD sensitivity. Instead, the high electrode firing is modulated by a lower rate which may be enhancing ITD sensitivity via a mechanism more akin to that described in [53]. In [53], “jittering” of the pulse timing of a high-rate electrical pulse train led to greater ITD sensitivity. Ultimately, more work is needed to understand how the TLE strategy is enabling improved binaural hearing in CI users as it may have implications on the choice of f_{lim} .

VI. Conclusion

This paper describes the TLE sound coding strategy, which has the potential for improving access to temporal fine structure cues in bilateral cochlear implant users. The effect of choice of lower limit of the TLE modulator on word recognition and binaural benefit was examined. Word recognition performance with TLE was comparable to that of the ACE sound coding strategy, as long as the lower limit of the modulator was 200 Hz. When listening in noise, binaural benefits were observed in individual listeners but appeared unrelated to changes in the lower limit of the modulator. Further work is needed to understand the underlying mechanism by which the TLE strategy is enabling binaural benefit.

Acknowledgment

The authors deeply appreciate the listeners who travelled to the University of Wisconsin-Madison to participate in this research. We would also like to thank Kay Khang for her assistance in collecting the data from normal-hearing listeners, Dr. Hussnain Ali and Prof. John Hansen for providing technical support in the programming of the CCI-MOBILE research platform, Shelly Godar who organized travel arrangements for the participants, and Prof. Ruth Litovsky for her mentorship and allowing these experiments to be conducted using her laboratory facilities.

This work was funded by the National Institutes of Health-National Institute on Deafness and Other Communication Disorders under Grant R03DC015321 to Kan, and in part by the National Institutes of Health-Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant U54HD090256 to the Waisman Center, and the National Natural Science Foundation of China Grant No. 11704129 and 61771320 and the Guangdong Basic and Applied Basic Research Roundation Grant No. 2020A1515010386 to Meng. (Corresponding author: Qinglin Meng)

Biography



Alan Kan (S'03-M'10) received the B.E. (Telecommunications) and Ph.D. in Engineering from the University of Sydney, Australia in 2002 and 2010, respectively.

From 2010 to 2019, he worked at the Waisman Center, University of Wisconsin-Madison, first as a Post-doctoral Research Associate and later promoted to Assistant Scientist. In 2019, he joined the School of Engineering at Macquarie University, Australia. His research interests include signal processing, audio virtual reality, psychoacoustics, binaural hearing, and cochlear implants.

Dr. Kan is an elected committee member of the Acoustical Society of America - Psychological and Physiological Acoustics Technical Committee, and Editorial Board Member of the Journal of Speech, Language and Hearing Research.



Qinglin Meng (M'20) received the B.E. (electrical engineering) from the Harbin Engineering University, China in 2008 and the Ph.D. in engineering from the Shanghai Acoustics Laboratory, Institute of Acoustics, Chinese Academy of Sciences in 2013.

From 2013 to 2016, he was with Shenzhen University as a post-doc. From 2017 to 2018, he visited Department of Biomedical Sciences, City University of Hong Kong for one year. In 2016, he joined the Acoustics Laboratory, School of Physics and Optoelectronics, South China University of Technology. His research interests are in cochlear implant signal processing, sound perception, and hearing health.

Dr. Meng is a member of the Acoustical Society of China - Acoustic Education Technical Committee and a member of the Chinese Computer Federation Task Force on Speech Dialogue and Auditory Processing.

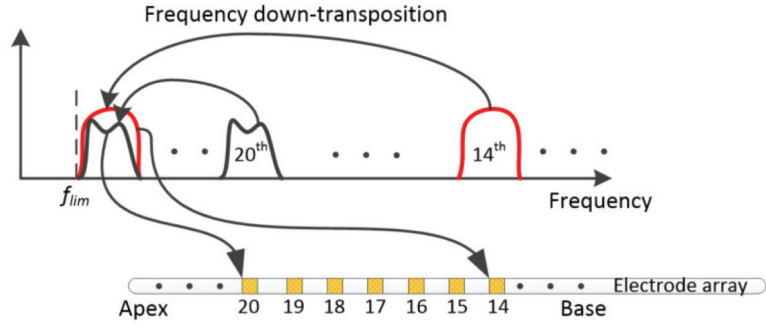
References

- [1]. Wouters J, McDermott HJ, and Francart T. "Sound Coding in Cochlear Implants: From electric pulses to hearing". In: *IEEE Signal Processing Magazine* 322 (3 2015), pp. 67–80.
- [2]. Wilson BS, Finley CC, Lawson DT, et al. "Better speech recognition with cochlear implants." In: *Nature* 3526332 (7 1991), pp. 236–8. [PubMed: 1857418]
- [3]. Wilson BS and Dorman MF. "The Surprising Performance of Present-Day Cochlear Implants". In: *IEEE Transactions on Biomedical Engineering* 546 (6 2007), pp. 969–972. [PubMed: 17554816]

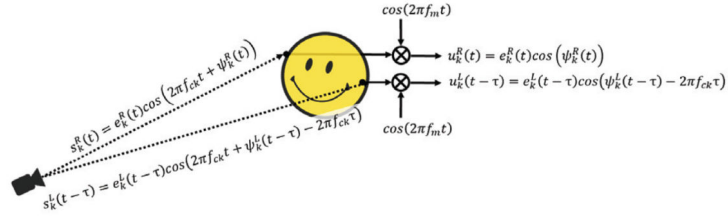
- [4]. Firszt JB, Holden LK, Skinner MW, et al. "Recognition of Speech Presented at Soft to Loud Levels by Adult Cochlear Implant Recipients of Three Cochlear Implant Systems". In: *Ear and Hearing* 254 (8 2004), pp. 375–387.
- [5]. Loizou PC, Hu Y, Litovsky R, et al. "Speech recognition by bilateral cochlear implant users in a cocktail-party setting." In: *The Journal of the Acoustical Society of America* 1251 (1 2009), pp. 372–83. [PubMed: 19173424]
- [6]. Litovsky RY, Parkinson A, and Arcaroli J. "Spatial Hearing and Speech Intelligibility in Bilateral Cochlear Implant Users". In: *Ear and Hearing* 304 (8 2009), pp. 419–431. [PubMed: 19455039]
- [7]. Dorman M, Yost W, Wilson B, et al. "Speech perception and sound localization by adults with bilateral cochlear implants". In: *Seminars in Hearing* 321 (2011), pp. 73–89.
- [8]. Macpherson EA and Middlebrooks JC. "Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited." In: *The Journal of the Acoustical Society of America* 1115 Pt 1 (5 2002), pp. 2219–2236. [PubMed: 12051442]
- [9]. Litovsky RY, Goupell MJ, Misurelli SM, et al. "Hearing with Cochlear Implants and Hearing Aids in Complex Auditory Scenes". In: *The Auditory System at the Cocktail Party*. Springer Handbook of Auditory Research, Vol 60 Ed. by Middlebrooks JC, Simon J, Popper A, et al. Cham, Switzerland: Springer, 2017 Chap. 10, pp. 261–291. ISBN: 978–3-319–51660-8.
- [10]. Aronoff J, Freed D, and Fisher L. "Cochlear implant patients' localization using interaural level differences exceeds that of untrained normal hearing listeners". In: *The Journal of the Acoustical Society of America* 1315 (2012), pp. 382–387.
- [11]. Litovsky RY, Goupell MJ, Godar S, et al. "Studies on bilateral cochlear implants at the University of Wisconsin's Binaural Hearing and Speech Laboratory." In: *Journal of the American Academy of Audiology* 236 (6 2012), pp. 476–94. [PubMed: 22668767]
- [12]. Kan A and Litovsky RY. "Binaural hearing with electrical stimulation". In: *Hearing Research* 322 (4 2015), pp. 127–137. [PubMed: 25193553]
- [13]. van Hoesel RJ. "Exploring the Benefits of Bilateral Cochlear Implants". In: *Audiology and Neurotology* 94 (2004), pp. 234–246. [PubMed: 15205551]
- [14]. Hochmair I, Nopp P, Jolly C, et al. "MED-EL Cochlear implants: state of the art and a glimpse into the future." In: *Trends in Amplification* 104 (12 2006), pp. 201–19. [PubMed: 17172548]
- [15]. Churchill TH, Kan A, Goupell MJ, et al. "Spatial hearing benefits demonstrated with presentation of acoustic temporal fine structure cues in bilateral cochlear implant listeners". In: *The Journal of the Acoustical Society of America* 1363 (9 2014), pp. 1246–1256. [PubMed: 25190398]
- [16]. Thakkar T, Kan A, Jones HG, et al. "Mixed stimulation rates to improve sensitivity of interaural timing differences in bilateral cochlear implant listeners". In: *The Journal of the Acoustical Society of America* 1433 (3 2018), pp. 1428–1440. [PubMed: 29604701]
- [17]. van Hoesel RJ, Jones GL, and Litovsky RY. "Interaural Time-Delay Sensitivity in Bilateral Cochlear Implant Users: Effects of Pulse Rate, Modulation Rate, and Place of Stimulation". In: *Journal of the Association for Research in Otolaryngology* 104 (12 2009), pp. 557–567. [PubMed: 19513792]
- [18]. van Hoesel RJM and Tyler RS. "Speech perception, localization, and lateralization with bilateral cochlear implants." In: *The Journal of the Acoustical Society of America* 1133 (3 2003), pp. 1617–30. [PubMed: 12656396]
- [19]. Zirn S, Arndt S, Aschendorff A, et al. "Perception of Interaural Phase Differences With Envelope and Fine Structure Coding Strategies in Bilateral Cochlear Implant Users." In: *Trends in Hearing* 20 (9 2016).
- [20]. Meng Q, Zheng N, and Li X. "A temporal limits encoder for cochlear implants". In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) IEEE*, 4 2015, pp. 5863–5867. ISBN: 978–1-4673–6997-8.
- [21]. Meng Q, Zheng N, and Li X. "Mandarin speech-in-noise and tone recognition using vocoder simulations of the temporal limits encoder for cochlear implants." In: *The Journal of the Acoustical Society of America* 1391 (1 2016), pp. 301–10. [PubMed: 26827026]
- [22]. Meng Q, Wang X, Zheng N, et al. "Binaural hearing measured with the temporal limits encoder using a vocoder simulation of cochlear implants". In: *Acoustical Science and Technology* 411 (1 2020), pp. 209–213.

- [23]. McKay CM, McDermott HJ, and Clark GM. "Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees". In: *The Journal of the Acoustical Society of America* 965 (11 1994), pp. 2664–2673. [PubMed: 7983272]
- [24]. Kan A and Meng Q. "Spatial Release From Masking in Bilateral Cochlear Implant Users listening to the Temporal Limits Encoder Strategy". In: *Proceedings of the 23rd International Congress on Acoustics Aachen, Germany, 2019*.
- [25]. Hansen JH, Ali H, Saba JN, et al. "CCi-MOBILE: Design and Evaluation of a Cochlear Implant and Hearing Aid Research Platform for Speech Scientists and Engineers". In: *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI) IEEE, 5 2019*, pp. 1–4. ISBN: 978-1-7281-0848-3.
- [26]. Ali H, Lobo AP, and Loizou PC. "Design and evaluation of a personal digital assistant-based research platform for cochlear implants." In: *IEEE Transactions on Biomedical Engineering* 6011 (11 2013), pp. 3060–73. [PubMed: 23674422]
- [27]. Kan A, Peng ZE, Moua K, et al. "A systematic assessment of a cochlear implant processor's ability to encode interaural time differences". In: *2018 AsiaPacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) Honolulu, Hawaii: IEEE, 11 2018*, pp. 382–387. ISBN: 978-9-8814-7685-2.
- [28]. Meng Q and Yu G. "The "temporal limits encoder" for cochlear implants: its potential advantages and an FFT-based algorithm". In: *Conference on Implantable Auditory Prostheses Lake Tahoe, CA, 2017*.
- [29]. Lehiste I and Peterson GE. "Linguistic Considerations in the Study of Speech Intelligibility". In: *The Journal of the Acoustical Society of America* 313 (3 1959), pp. 280–286.
- [30]. IEEE. "IEEE recommended practice for speech quality measurements". In: *IEEE Transactions on Audio and Electroacoustics* 173 (1969).
- [31]. Levitt H. "Transformed up-down methods in psychoacoustics." In: *The Journal of the Acoustical Society of America* 492 (2 1971), Suppl 2:467+.
- [32]. Schütt HH, Harmeling S, Macke JH, et al. "Painfree " and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data." In: *Vision research* 122 (5 2016), pp. 105–123. [PubMed: 27013261]
- [33]. Robinson JD, Baer T, and Moore BCJ. "Using transposition to improve consonant discrimination and detection for listeners with severe high-frequency hearing loss." In: *International journal of audiology* 466 (6 2007), pp. 293–308. [PubMed: 17530514]
- [34]. Loizou PC, Poroy O, and Dorman M. "The effect of parametric variations of cochlear implant processors on speech understanding." In: *The Journal of the Acoustical Society of America* 1082 (8 2000), pp. 790–802. [PubMed: 10955646]
- [35]. Nadol JB. "Patterns of neural degeneration in the human cochlea and auditory nerve: implications for cochlear implantation." In: *Otolaryngology-head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 1173 Pt 1 (9 1997), pp. 220–8. [PubMed: 9334769]
- [36]. Kawano A, Seldon HL, Clark GM, et al. "Intracochlear factors contributing to psychophysical percepts following cochlear implantation." In: *Acta otolaryngologica* 1183 (6 1998), pp. 313–26.
- [37]. Moore BC, Huss M, Vickers DA, et al. "A test for the diagnosis of dead regions in the cochlea." In: *British journal of audiology* 344 (8 2000), pp. 205–24. [PubMed: 10997450]
- [38]. Arenberg Bierer J. "Probing the Electro-Neuron Interface With Focused Cochlear Implant Stimulation". In: *Trends in Amplification* 142 (6 2010), pp. 84–95. [PubMed: 20724356]
- [39]. Garadat SN, Litovsky RY, Yu G, et al. "Effects of simulated spectral holes on speech intelligibility and spatial release from masking under binaural and monaural listening." In: *J Acoust Soc Am* 1272 (2 2010), pp. 977–989. [PubMed: 20136220]
- [40]. Goupell MJ, Kan A, and Litovsky RY. "Spatial attention in bilateral cochlear-implant users." In: *The Journal of the Acoustical Society of America* 1403 (9 2016), pp. 1652–1662. [PubMed: 27914414]
- [41]. Goupell MJ, Stakhovskaya OA, and Bernstein JGW. "Contralateral Interference Caused by Binaurally Presented Competing Speech in Adult Bilateral Cochlear-Implant Users." In: *Ear and hearing* 391 (8 2018), pp. 110–123. [PubMed: 28787316]

- [42]. Ketten DR, Skinner MW, Wang G, et al. "In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays." In: *The Annals of otology, rhinology & laryngology*. Supplement 175 (11 1998), pp. 1–16.
- [43]. Landsberger DM, Svrakic M, Roland JT, et al. "The relationship between insertion angles, default frequency allocations, and spiral ganglion place pitch in cochlear implants." In: *Ear and hearing* 365 (2015), pp. 207–13.
- [44]. Kan A, Stoelb C, Litovsky RY, et al. "Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear-implant users." In: *The Journal of the Acoustical Society of America* 1344 (10 2013), pp. 2923–36. [PubMed: 24116428]
- [45]. Kan A, Litovsky RY, and Goupell MJ. "Effects of interaural pitch matching and auditory image centering on binaural sensitivity in cochlear implant users." In: *Ear and hearing* 363 (1 2015), pp. 62–8.
- [46]. Kan A, Goupell MJ, and Litovsky RY. "Effect of channel separation and interaural mismatch on fusion and lateralization in normal-hearing and cochlear-implant listeners". In: *The Journal of the Acoustical Society of America* 1462 (8 2019), pp. 1448–1463. [PubMed: 31472555]
- [47]. Yoon Y-S, Liu A, and Fu Q-J. "Binaural benefit for speech recognition with spectral mismatch across ears in simulated electric hearing." In: *The Journal of the Acoustical Society of America* 1302 (8 2011), pp. 94–100.
- [48]. Yoon Y-S, Shin Y-R, and Fu Q-J. "Binaural benefit with and without a bilateral spectral mismatch in acoustic simulations of cochlear implant processing." In: *Ear and hearing* 343 (2013), pp. 273–9. [PubMed: 22968427]
- [49]. Goupell MJ, Stoelb CA, Kan A, et al. "The Effect of Simulated Interaural Frequency Mismatch on Speech Understanding and Spatial Release From Masking." In: *Ear and hearing* (1 2018).
- [50]. Hall JW, Haggard MP, and Fernandes MA. "Detection in noise by spectro-temporal pattern analysis." In: *The Journal of the Acoustical Society of America* 761 (7 1984), pp. 50–6. [PubMed: 6747111]
- [51]. Nie K, Stickney G, and Zeng F-G. "Encoding frequency modulation to improve cochlear implant performance in noise." In: *IEEE transactions on bio-medical engineering* 521 (1 2005), pp. 64–73. [PubMed: 15651565]
- [52]. Smith ZM, Delgutte B, and Oxenham AJ. "Chimaeric sounds reveal dichotomies in auditory perception." In: *Nature* 4166876 (3 2002), pp. 87–90. [PubMed: 11882898]
- [53]. Laback B and Majdak P. "Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates." In: *Proceedings of the National Academy of Sciences of the United States of America* 1052 (1 2008), pp. 814–7. [PubMed: 18182489]



(a)



(b)

Figure 1: TLE processing. (a) illustrates TLE frequency down-transposition of channel information and (b) is a schematic showing that ITD information is encoded when the TLE strategy applied to both ears.

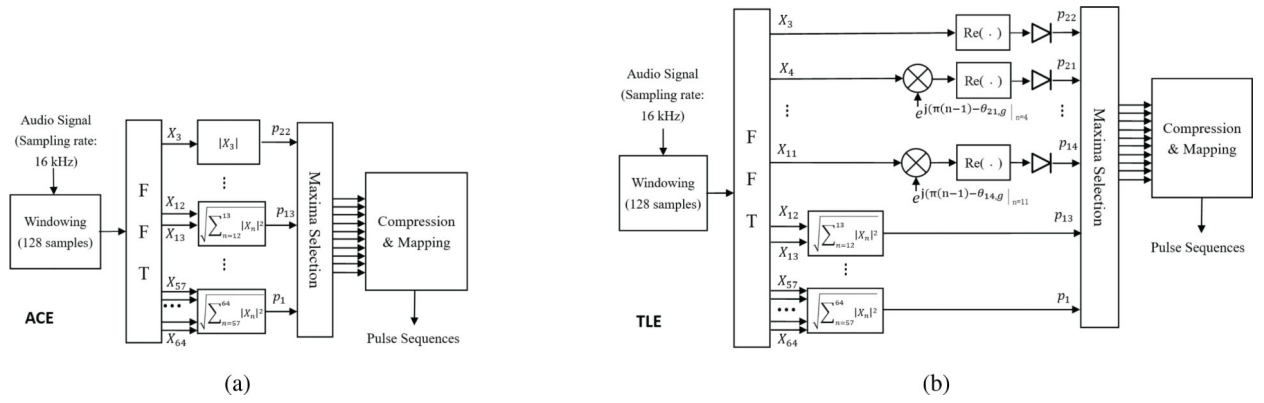


Figure 2: The ACE and TLE implementations on the CCI-MOBILE used in the experiment.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

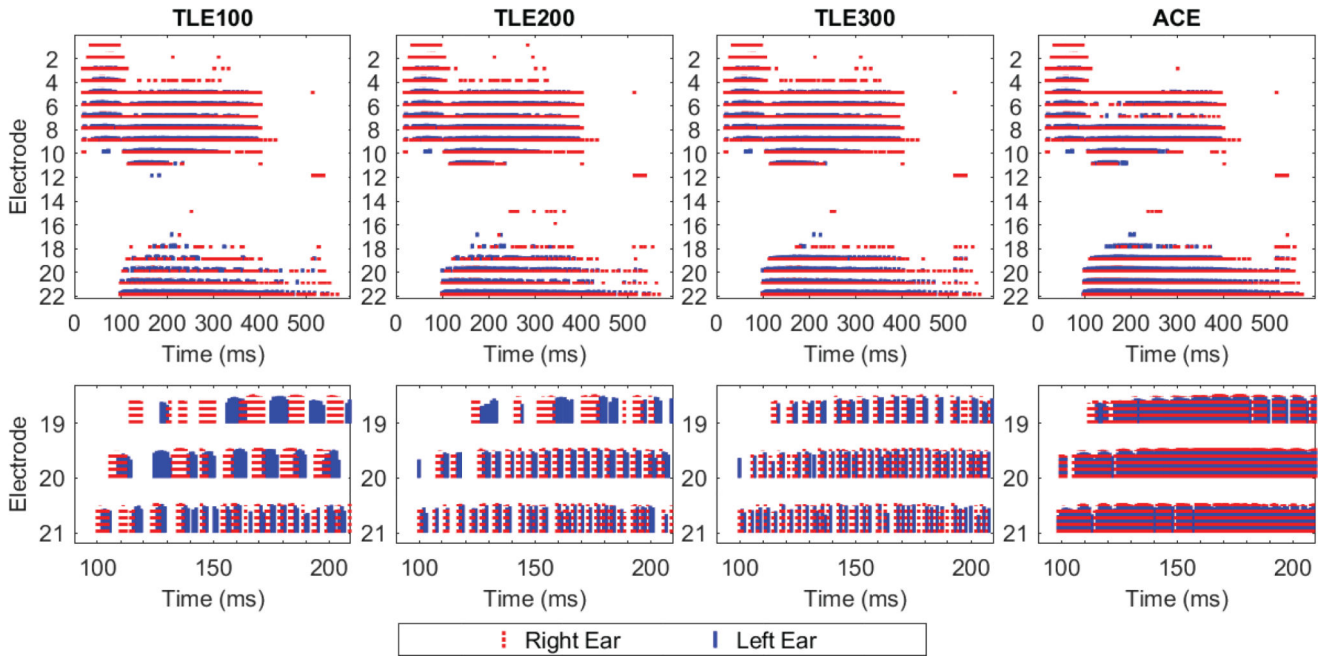


Figure 3: Electrode output for the different strategy conditions tested. Each red and blue line represents a single electric pulse that is sent out of a particular electrode. The input stimulus was the word “Jane” presented towards the right ear. The top panels show the electrode output on all channels, and the bottom panels show an enlarged view of some of the low frequency channels where the TLE processing has been applied. In the Cochlear system, low frequency electrodes have higher channel numbers.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

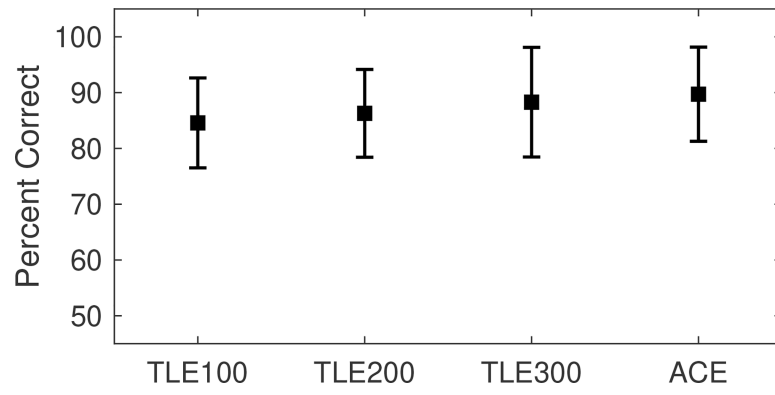


Figure 4: Word recognition performance in quiet. Error bars show the standard deviation.

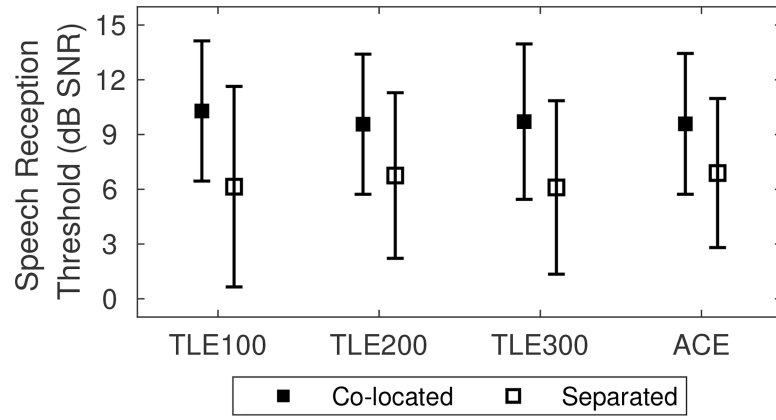


Figure 5: Speech reception threshold (SRT) for Co-located and Separated conditions.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

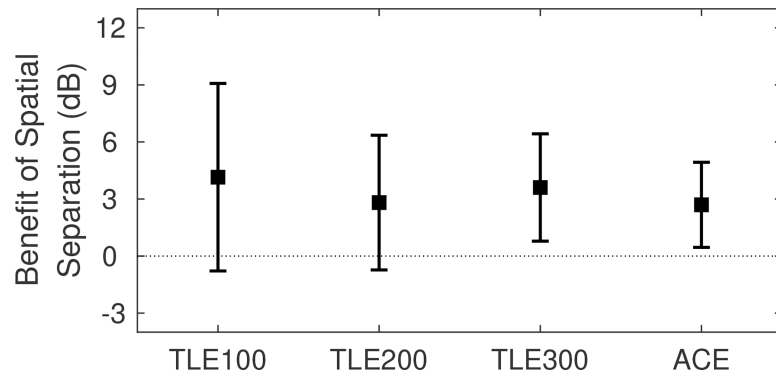


Figure 6:
Spatial Release from Masking (SRM).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

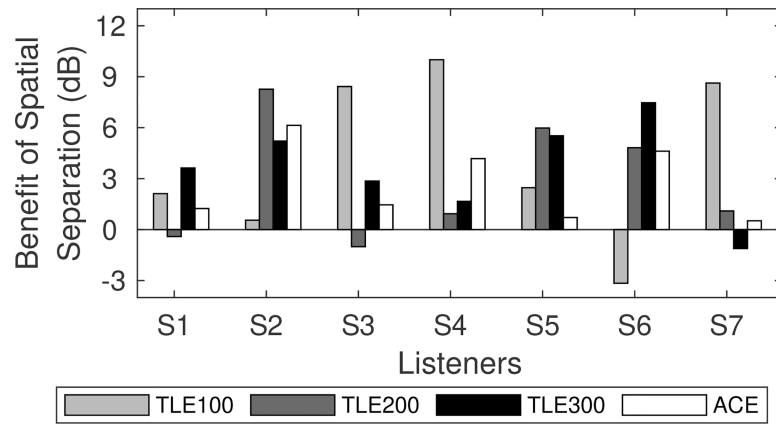


Figure 7:
Binaural benefit per listener.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table I:

Individual percent correct word recognition in quiet

ID	TLE100 (%)	TLE200 (%)	TLE300 (%)	ACE (%)
S1	90	92	96	90
S2	92	94	94	98
S3	88	92	92	94
S4	92	92	98	100
S5	82	78	86	88
S6	72	80	70	78
S7	76	76	82	80
MEAN	84.6	86.3	88.3	89.7

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table II:

Individual speech reception thresholds for word recognition in noise

Co-located Condition				
ID	TLE100 (dB)	TLE200 (dB)	TLE300 (dB)	ACE (dB)
S1	8.3	9.0	9.8	6.6
S2	4.7	5.9	4.6	5.5
S3	8.1	7.3	7.2	8.2
S4	10.1	5.6	7.8	11.4
S5	12.2	12.7	8.2	11.5
S6	12.0	16.2	17.5	16.6
S7	16.7	10.3	12.7	7.3
MEAN	10.30	9.57	9.69	9.59
Separated Condition				
ID	TLE100 (dB)	TLE200 (dB)	TLE300 (dB)	ACE (dB)
S1	6.1	9.5	6.2	5.4
S2	4.1	-2.3	-0.6	-0.6
S3	-0.4	8.3	4.4	6.8
S4	0.1	4.6	6.1	7.2
S5	9.7	6.7	2.7	10.8
S6	15.2	11.4	10.1	12.0
S7	8.1	9.2	13.9	6.7
MEAN	6.13	6.77	6.11	6.90

Table III:

TLE setting yielding greatest binaural benefit

ID	f_{im}	SRM_{TLE} (dB)	SRM_{ACE} (dB)	Quiet_{TLE-ACE} (%)
S1	300	3.6	1.2	6
S2	200	8.3	6.1	-4
S3	100	8.1	1.5	-6
S4	100	10.0	4.2	-8
S5	200	6.0	0.7	-10
S6	300	7.4	4.6	-8
S7	100	8.6	0.5	-4
MEAN		7.4	2.7	-4.9

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript