A systematic assessment of a cochlear implant processor's ability to encode interaural time differences

Alan Kan*, Z. Ellen Peng, Keng Moua and Ruth Y. Litovsky
University of Wisconsin-Madison
*E-mail: ahkan@waisman.wisc.edu Tel: +1-608-262-7483

***Abstract*—Bilateral cochlear implantation is becoming the standard of care for patients with sensorineural hearing loss with demonstrated improvements over unilateral use in everyday tasks, such as sound localization ability. However, even with bilateral implantation, performance in these tasks is still poorer than that of normal hearing listeners. The gap in performance has often been attributed to the poor encoding of fine structure interaural time differences (ITDs) by clinical processor. However, in theory, the signal processing employed in clinical processors should still encode envelope ITDs with some degree of fidelity. In this work, we quantitatively measured the ability of Cochlear CP910 processors to encode envelope ITDs, while running the Advanced Combinational Encoder (ACE) strategy. Results suggest that while the processors are able to support relatively good envelope encoding, the peak-picking approach of the ACE strategy degrades the computation of ITDs by encoding spectral information in different frequency regions in the two ears. Our results may explain the poorer sound localization performance observed in cochlear implant users who use the ACE strategy, but cannot account for the poorer sound localization performance observed in cochlear implant users in general.**

I. Introduction

Cochlear implants (CIs) have been remarkably successful in restoring hearing abilities to patients who are profoundly deaf. CIs work by taking an incoming sound and converting it into pulsatile electrical stimulation. While there are some signal processing differences between manufacturers, all CIs typically follow the same signal processing architecture. In general, the incoming sound is band-passed filtered into a small number of channels (12 to 22, depending on manufacturer). Then, the signal envelope in each channel is extracted, compressed, and used to modulate the amplitudes of an electrical pulse train, typically at a rate of ≥900 pulses per second (pps) per channel. These pulse trains are allocated to different electrodes which stimulate different parts of the cochlea. By taking advantage of the place-to-frequency organization of the cochlea, the CI is able to transmit the frequencies of the incoming sound to the brain, albeit with poorer spectral resolution than normal hearing (NH).

Although CIs were originally designed for unilateral use, bilateral implantation (i.e., having a CI in each ear) has become more common in recent years [1], [2]. When listening with two CIs compared with one CI, most patients demonstrate improvements in sound localization abilities and speech-in-noise understanding [3], [4]. However, there is still a gap in performance between bilateral CI users and NH listeners on these tasks [5]–[7]. The gap in performance in sound localization tasks is often attributed to the diminished access to the interaural timing difference (ITD), an important low-frequency cue that is used by NH listeners when locating sounds [8], [9]. ITD cues arise because a sound off the midline arrives at different times at the two ears, and this difference is assumed to be computed in the brainstem in a frequency-dependent manner to determine the location of the sound [10], [11].

There are many reasons why bilateral CI users do not appear to use ITDs for sound localization (see [12] for review). One reason that is commonly cited is that CI processors usually only encode the signal envelope, while the temporal fine structure (TFS) of the acoustic signal is typically discarded. Hence, CIs do not encode TFS ITDs in the electrical signal [13]–[15]. However, in the NH system, ITDs can also be computed from the signal envelope. In theory, envelope ITDs should still be encoded by clinical processors, but this has not been thoroughly investigated. One important factor that may affect the fidelity of envelope ITD encoding is the independent sampling clocks in the processors at the two ears. This may lead to an incoming sound being sampled at different phases in each ear, which may introduce spurious envelope ITD cues.

In this work, we quantitatively assessed whether clinical processors can encode envelope ITD cues with some degree of fidelity. As a first step, we focused on Cochlear Ltd's Nucleus processors because they are widely used [16]. In particular, we focus on the Advance Combinational Encoder (ACE) sound coding strategy [17]–[19], which is the default signal processing strategy used nowadays in Cochlear devices. The ACE strategy uses a Fast Fourier Transform (FFT) to separate the incoming acoustic signal into 22 band-limited channels. The strategy is unique among the different CI manufacturers in that it does not attempt to encode the entire frequency spectrum of the incoming acoustic signal at once. Instead, it employs a peak-picking technique aimed at transmitting information from

This work was supported by grants from the National Institute of Health-National Institute on Deafness and Other Communication Disorders (NIH-NIDCD R03-DC015321 to AK and R01-DC003083 to RYL) and the National Institute of Health-Eunice Kennedy Shriver National Institute on Child Health and Human Development (NIH-NICHD U54-HD090256).

the N highest peaks (typically 8) out of 22 in any time frame. The reduction in the number of frequency peaks that are encoded is done to minimize inter-channel crosstalk due to spread of current, which has been shown to have a negative impact on speech understanding [20], [21]. However, this peak-picking technique may distort the envelope transmitted in each frequency channel which may lead to poor envelope ITD encoding. In addition, peak-picking is independent in the two ears, which is likely to reduce the probability of having same-frequency channels in the two ears picked at the same time. This will limit the brain's ability to compute the ITD from the envelope within a frequency band because the signal envelope may only be available in one ear.

Only one previous attempt has been made to quantitatively measure the ACE strategy's ability to encode ITDs [22]. In that work, ITDs encoded by Cochlear CP810 (N5) and Freedom processors were measured in the free field on a KEMAR manikin for different sound source locations along the horizontal plane. A range of stimuli were used for the measurements including a 2-kHz pure tone, 2-kHz sinusoidal amplitude-modulated tone with 10 and 100-Hz modulation frequency, white noise, and speech. Results showed that ITDs extracted from the electrical stimulation patterns had large deviations from the expected ITD (up to 2.5 ms). However, no systematic pattern of errors were observed as a function of sound source direction or stimulus, which makes it difficult to understand the source of the ITD encoding errors.

In the current work, we took a more systematic approach in quantitatively assessing a clinical processor's ability to encode envelope ITDs. First, we assessed the impact of sampling and processing on envelope encoding. Independent sampling of the signal waveform at the two ears may lead to jitter in the envelope ITD which may render envelope ITDs unreliable. Second, we assessed the ability of a pair of clinical processors to encode real ITDs by applying head-related transfer functions (HRTFs) to the acoustic stimulus. HRTFs describe the direction-dependent acoustical transformation of a sound from its source location to the listener's ears which include ITDs. These measurements allow us to determine the effects of physical acoustics on how well the ACE strategy can encode real-world ITD cues.

II. General Methods

All measurements were performed using the "direct connect" method by providing stimuli via an audio input cable to the accessory port of a Cochlear CP910 (marketed as N6) processor. Measurement using the direct connect method is superior to free field measurements previously used in [22] because it minimizes systematic bias such as room reflections and loudspeaker transfer functions. It allows us to directly compare ITDs computed from the electrode stimulation output with ITDs computed from the input stimulus waveform. All stimuli were generated using MATLAB and presented to the processor via a Tucker Davis Technologies (TDT) System3 equipped with RP2.1, HB7 and PA5 units. The electrical stimulation output from the CP910 was recorded using a National Instruments USB-6343 data acquisition card (NIDAQ) connected to an Implant-in-a-box containing a CI24R electrode array. The NIDAQ has an overall maximum sampling rate of 500000 Hz which is shared across the 32 recording channels. Hence, the actual sampling rate is dependent on the number of active recording channels. In this work, we made recordings on one channel alone, and twenty-two channels simultaneously, resulting in sampling rates of 500000 and 22727 Hz per channel, respectively.

Cochlear's Custom Sound software (Version 5.0) was used to create a map for the CP910 processor. A map defines the signal processing and processor parameters. The map used for these measurements had threshold stimulation levels set to 100 clinical units (CU) on all channels, and a dynamic range of 150 CUs. Biphasic monopolar stimulation was used with a pulse duration of 25 μs/phase and MP1+2 ground configuration. Stimulation rates on each channel was 900 pps and the number of channels to stimulate per cycle (Maxima) was set to 8. The default frequency allocation table (table number 22 in Custom Sound) was used, which allocates frequencies up to about 8000 Hz to all electrodes. Table 1 shows the frequency allocations for each channel. Front end processing (i.e., SCAN, ADRO and ASC[1]) and the telecoil accessory were switched off, and the accessory port mixing ratio was set to accessory only (disables microphone input). All other map parameters were left at their default settings.

TABLE I
FREQUENCY ALLOCATION.

Channel Number	Lower Frequency (Hz)	Upper Frequency (Hz)	Center Frequency (Hz)
1	6938	7938	7438
2	6063	6938	6500.5
3	5313	6063	5688
4	4688	5313	5000.5
5	4063	4688	4375.5
6	3563	4063	3815
7	3063	3563	3313
8	2688	3063	2875.5
9	2313	2688	2500.5
10	2063	2313	2188
11	1813	2063	1938
12	1563	1813	1688
13	1313	1563	1438
14	1188	1313	1250.5
15	1063	1188	1125.5
16	938	1063	1000.5
17	813	938	875.5
18	688	813	750.5
19	563	688	625.5
20	438	563	500.5
21	313	438	375.5
22	188	313	250.5

[1] SCAN is a sound scene classification algorithm; ADRO (Adaptive Dynamic Range Optimziation) is an amplification strategy that optimizes gain settings in each frequency channel by accounting for a user's audible dynamic range; ASC (Autosensitivity Control) is an automatic microphone sensitivity control algorithm.

The target stimuli for the recordings comprised of transposed tones with a 30-Hz modulation rate. The carrier frequencies of the transposed tones matched the center frequencies listed in Table I; that is, one transposed tone was created per channel. All stimuli were 300 ms long and were initially created at a sampling rate of 44100 Hz but was later resampled in MATLAB to 48828 Hz to match the playback sampling rate of the TDT System.

For analysis, the envelope of the electrical stimulation pattern on each channel was extracted from the recording. The envelope was calculated by taking the Hilbert transform of the recorded signal and low-pass filtering the magnitude of the output of the Hilbert transform with a fourth-order Butterworth filter with a cutoff frequency at 50 Hz.

III. Measurement 1 – Effect of Sampling

A. Methods

The aim of this measurement was to assess the impact of sampling and processing on envelope encoding by measuring the variations in the electrical stimulation output. The results allow us to infer the best case envelope ITD encoding of the CP910 processor running the ACE strategy. Fifty, single-channel recordings were made of a stimulus that consisted of a 100 ms synchronization sine tone followed by 200 ms silence followed by a transposed tone (see Fig. 1). The frequency of the sine tone and the carrier frequency of the transposed tone matched the corresponding channel listed in Table I.

Recordings were analyzed by calculating the temporal differences in the signal envelope onset between recordings. The recordings were first time-aligned by cross-correlation of the envelope derived from the 100 ms sine tone component of the recording stimulus. Two metrics were then calculated from the envelope of the transposed component: (1) Mean onset jitter – defined as the difference in time between pairs of recordings for the first rise in the envelope in the transposed component of the signal. The time difference was calculated at the 50% point relative to peak amplitude; (2) Mean waveform jitter – calculated by cross-correlating the envelope of the transposed component of the signal.

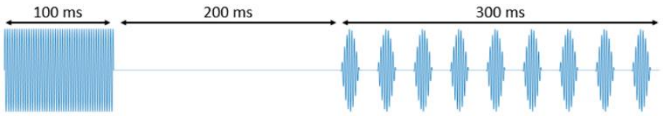

Figure 1 shows an example of the recording signal. Each recording signal consisted of a 100 ms synchronization tone, 200 ms silence, followed by the stimulus of interest (transposed tone or noise).

B. Results

Figure 2 shows example recordings made from a single electrode stimulated through the CP910 sound processor. It can be seen that the CP910 is capable of encoding the envelope of the target stimulus with a relatively high degree of fidelity. Mean onset and waveform jitter derived from the recordings are listed in Table II. The onset jitter was typically at least double that of the waveform jitter. Waveform jitter was typically larger towards the ends of the electrode array compared with the channels in the middle of the array.

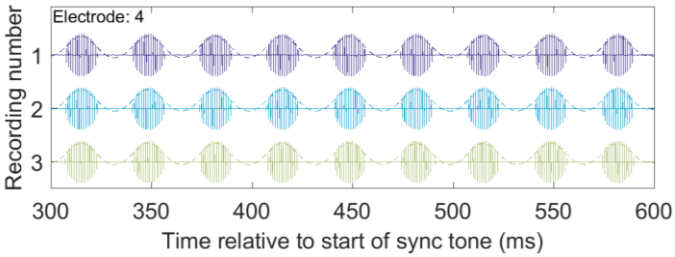

Figure 2 shows examples of the electric stimulation output recorded on electrode number 4. The dashed line shows the extracted envelope.

TABLE II
ESTIMATED ENVELOPE JITTER OF CP910 PROCESSORS.
STANDARD DEVIATION IS SHOWN IN PARENTHESIS.

Channel Number	Mean Onset Jitter (μs)	Mean Waveform Jitter (μs)
1	261.7 (179.1)	161.6 (125.3)
2	284.4 (146.9)	201.7 (124.8)
3	155.0 (48.9)	46.0 (27.6)
4	72.9 (38.6)	29.9 (20.6)
5	113.9 (45.8)	35.1 (23.9)
6	43.6 (24.9)	20.7 (13.1)
7	68.1 (33.4)	20.2 (12.1)
8	52.0 (45.4)	24.3 (15.2)
9	80.0 (54.1)	47.9 (29.7)
10	87.4 (35.7)	23.4 (17.8)
11	54.5 (33.8)	20.2 (18.9)
12	59.8 (34.4)	21.4 (13.6)
13	106.6 (58.5)	44.5 (31.4)
14	91.8 (40.3)	27.7 (19.0)
15	105.0 (47.7)	27.3 (22.0)
16	114.3 (46.3)	28.8 (19.8)
17	108.5 (48.5)	53.4 (26.0)
18	94.5 (42.0)	39.2 (30.9)
19	151.0 (61.9)	36.2 (27.8)
20	154.4 (52.8)	43.6 (28.6)
21	192.6 (67.1)	105.9 (59.3)
22	242.6 (129.0)	72.1 (73.8)
MEAN	122.5 (71.1)	51.4 (47.5)

C. Discussion

The relatively good envelope encoding found from these measurements suggest that the CP910 processors are able to encode envelope ITDs with some degree of fidelity. Two metrics were used to quantify the temporal precision of envelope encoding. Jitter in the first onset of the envelope appeared to be relatively large (~120 μs on average), especially if one considers that the range of human ITDs is typically within 700 μs (see Table III, input ITD columns). However, when the timing of the whole waveform is taken into account, temporal jitter was typically much better (~50 μs). Overall, both types of jitter appear to be within the just noticeable ITD observed in bilateral CI users (typically greater than 100 μs) when measured with a synchronized bilateral direct

TABLE III
ITDS FOR EACH CHANNEL DERIVED FROM HEAD-RELATED TRANSFER FUNCTIONS (INPUT) AND ELECTRIC STIMULATION RECORDINGS FROM CP910 PROCESSORS (OUTPUT). NEGATIVE ITDS MEAN THE LEFT SIGNAL LEADS THE RIGHT, AND VICE VERSA. STANDARD DEVIATION IS SHOWN IN PARENTHESIS.

Channel	Azimuth									
	0°		20°		40°		60°		80°	
	Input ITD (μs)	Output ITD (μs)	Input ITD (μs)	Output ITD (μs)	Input ITD (μs)	Output ITD (μs)	Input ITD (μs)	Output ITD (μs)	Input ITD (μs)	Output ITD (μs)
19	0	83.8 (42.4)	0	69.0 (103.7)	181.4	426.8 (165.8)	294.8	554.8 (176.3)	362.8	605.5 (154.7)
15	0	82.4 (45.0)	0	-125.4 (59.6)	0	125.2 (90.1)	362.8	603.1 (82.7)	544.2	761.3 (64.1)
11	0	20.2 (35.8)	45.4	330.7 (77.2)	90.7	287.2 (89.4)	476.2	715.1 (101.7)	453.5	642.2 (85.4)
7	0	-49.8 (17.0)	113.4	202.0 (16.4)	90.7	181.1 (20.7)	362.8	466.3 (16.8)	453.5	604.6 (21.7)
3	0	107.6 (36.3)	181.4	419.2 (34.0)	22.7	266.4 (41.8)	90.7	292.7 (42.6)	340.1	557.5 (51.6)
MEAN	0	48.8 (36.6)	68.0	179.1 (65.8)	77.1	257.3 (95.7)	317.5	526.4 (100.4)	430.8	634.2 (87.7)

stimulation platform [23]–[25]. Jitter appeared to be larger at the ends of the array which is likely due to edge effects of the FFT processing. Considering that psychophysical listening tests have shown that bilateral CI users rely less on the onset compared with the peak of the signal envelope for ITD judgments [26], [27], it would appear that the poor sensitivity to ITDs reported when listening with clinical processors is unlikely due to the independent sampling and processing of the acoustic signal at the two ears. In the next set of measurements, we assess the effect of frequency-dependent attenuation introduced by the head and body on envelope ITD encoding with CP910 processors.

IV. MEASUREMENT 2 – EFFECT OF HEAD AND BODY

A. Methods

These measurements assess the ability of a pair of clinical processors running the ACE strategy to encode real-world ITDs obtained using measurements that mimic free field listening conditions. Real-world ITDs were introduced by filtering the stimulus with an HRTF for a particular sound source direction prior to presentation to a pair of CP910 processors programmed with the same map.

HRTFs were recorded on a KEMAR manikin in a single-walled soundproof booth. The booth had additional sound absorbing foam attached to the inside walls to reduce reflections. The booth houses a semi-circular array of loudspeakers (Cambridge SoundWorks) spaced 5° apart along the horizontal plane. The manikin was located in the center of the array at a distance of 1.2 m from the loudspeakers. HRTFs were measured using a blocked-ear technique [28] using Golay codes as the recording stimulus. HRTFs were derived from the recordings using standard techniques [29], with the recording systems transfer function deconvolved from the recordings using a pseudoinverse method [30]. For the current set of measurements, the HRTFs for loudspeaker positions from 0° to 80° (front to right) at 20° increments were used.

Due to the limited number of channels and sampling rate of our NIDAQ system, we were unable to assess the fidelity of envelope ITD encoding when all channels are simultaneously presented with an acoustic input. Hence, two sets of recordings were made. First, we assess fidelity of envelope ITD encoding at a single-channel level using the methodology of Measurement 1 (see Section III.A). Recordings were made on a few select channels that span the length of the electrode array (19, 15, 11, 7, and 3). The stimulus was a transposed tone with center frequency targeting the specific channel filtered by an HRTF, and then sent to two CP910 processors. This measurement mimics free field listening conditions of a narrowband stimulus. The input and output ITD was estimated by calculating the envelope of the input stimulus and recorded electrical stimulation output, respectively, and cross-correlating the left-side and right-side signals. Second, we measured the impact of the peak-picking processing of the ACE strategy on an HRTF-filtered transposed tone complex created by in-phase summation of transposed tones created for all electrode channel. This set of recordings allows us to understand how the CI processor encodes wideband spectral information into electrical stimulation, and how this process affects envelope ITD encoding.

B. Results

Input and output ITDs computed from single-channel recordings are shown in Table III for loudspeaker positions from 0° to 80° (front to right) at 20° increments. Positive and negative numbers imply right-side and left-side leading ITDs, respectively. The mean output ITD across 50 trials on each channel are typically much larger than the input ITD. On some occasions, when the source location is near 0° (front), the output ITD may point in the wrong direction in some channels (for example, channel 7 and 15 for source locations 0° and 20°, respectively), though averaging across channels still places the output ITD in the correct direction except for the 0° azimuth location.

Figure 3 shows the electrical stimulation pattern recorded from the right and left ear processors for the HRTF-filtered transposed tone complex. It can be seen that when the sound is at 0° (front), the peak-picking approach of the ACE strategy preferentially activates channel numbers between 10 and 2 which correspond to the frequency range of 2 – 7 kHz. As the

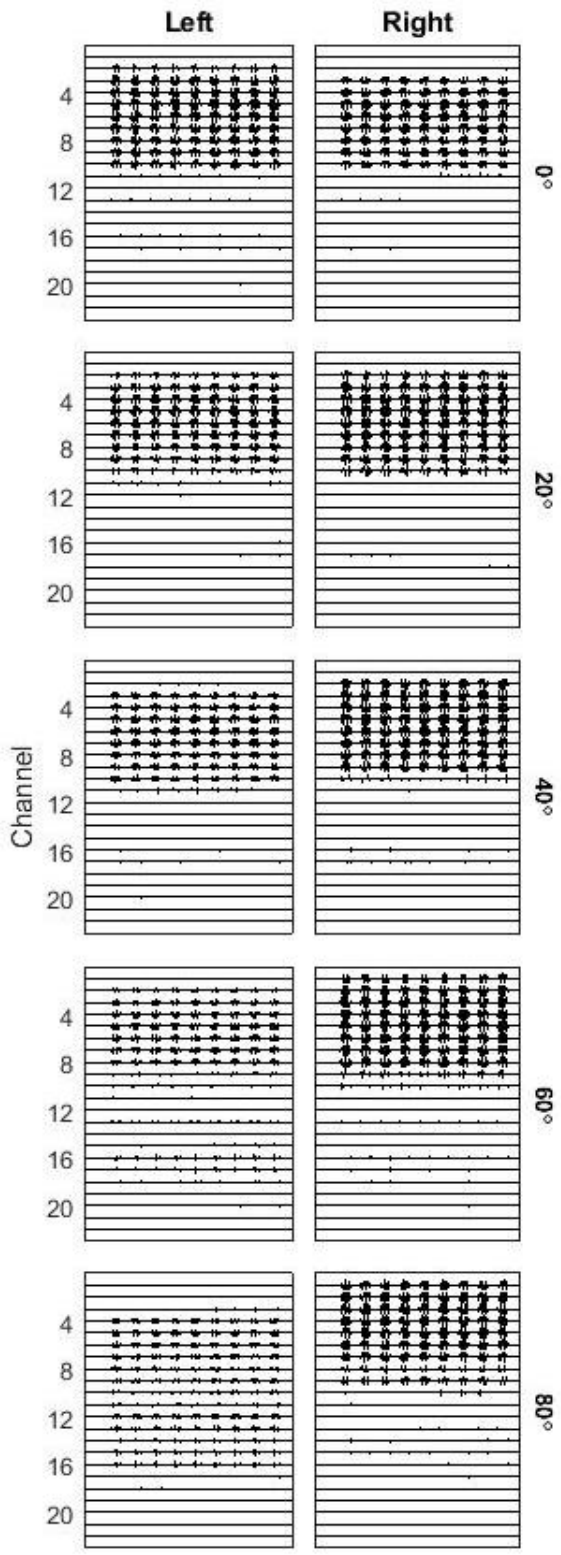

Figure 3 shows the electric stimulation output recorded from the left and right processors for a transposed-tone complex after it has been filtered by HRTFs. Each row of subplots is for a sound source direction and the columns are for the left and right ears, respectively. The horizontal lines in each subplot is the recording for one channel and the height of the line shows the amount of current that is being presented in that channel.

sound source moves towards the right, the active channels in the left and right ears appear to systematically shift lower and higher in frequency, respectively.

C. Discussion

At a single channel level, the ITDs encoded by the ACE strategy appear to be larger than that found in the input acoustic signal after HRTF filtering. While this may lead to errors in localization of a sound source, there is still a systematic trend in the ITDs such that larger ITDs are associated with more lateral positions. If the variance due to sampling (i.e., the standard deviation in each channel) can be reduced, it may be possible that CI users might be able to learn to map the larger ITDs to their corresponding lateral sound source positions.

Notably, the ACE strategy appears to degrade the usefulness of real-world ITDs in two ways. First, the active channels are typically in the higher frequencies which is the frequency region where interaural level, rather than time, differences are more dominant for sound localization tasks in NH listeners [31]. The bias towards higher frequencies is likely due to pre-emphasis filtering that is typically applied in CIs to enhance the higher frequencies [32], [33] coupled with the peak-picking approach of the ACE strategy. Second, the number of overlapping channels between the two processors decrease for sounds located off the midline. This limits the number of overlapping channels in the two ears for computing a consistent ITD. Considering timing differences between the ears only arise because sounds are located off the midline, this would present a problem for using ITDs to locate the direction of sounds.

V. Conclusions and Future Work

Our measurements suggest that for the stimuli used here, CI processors are able to encode ITD cues with some degree of fidelity despite the independent sampling at the two ears. However, it appears that peak-picking strategies, such as ACE, degrade the usefulness of ITDs because the full spectrum of the incoming sound is not encoded consistently in the same set of electrodes between the two ears. While our present results may explain the lack of access to ITDs by bilateral CI users who use the ACE strategy, it does not fully account for the poorer performance observed in CI users who use devices and sound processing strategies from other manufacturers. Hence, further work is need to understand how the strategies of other manufacturers may be degrading the encoding of ITD cues with electrical stimulation.

Acknowledgment

The authors would like to thank Jake Bergal for help with some of the initial measurements, and Zachary Smith and Aaron Parkinson from Cochlear Ltd for providing hardware and technical support.

References

[1] B. R. Peters, J. Wyss, and M. Manrique, "Worldwide trends in bilateral cochlear implantation.," *Laryngoscope*, vol. 120 Suppl, no. 5, pp. S17-44, May 2010.

[2] R. E. S. Lovett, D. A. Vickers, and A. Q. Summerfield, "Bilateral cochlear implantation for hearing-impaired children: Criterion of candidacy derived from an observational study," *Ear Hear.*, vol. 36, no. 1, pp. 14–23, 2015.

[3] R. Y. Litovsky, A. Parkinson, and J. Arcaroli, "Spatial Hearing and Speech Intelligibility in Bilateral Cochlear Implant Users," *Ear Hear.*, vol. 30, no. 4, pp. 419–431, Aug. 2009.

[4] R. S. Tyler, C. C. Dunn, S. A. Witt, and W. G. Noble, "Speech perception and localization with adults with bilateral sequential cochlear implants.," *Ear Hear.*, vol. 28, no. 2 Suppl, p. 86S–90S, Apr. 2007.

[5] H. Jones, A. Kan, and R. Y. Litovsky, "Comparing Sound Localization Deficits in Bilateral Cochlear-Implant Users and Vocoder Simulations With Normal-Hearing Listeners," *Trends Hear.*, vol. 18, Nov. 2014.

[6] P. C. Loizou, Y. Hu, R. Litovsky, G. Yu, R. Peters, J. Lake, and P. Roland, "Speech recognition by bilateral cochlear implant users in a cocktail-party setting.," *J. Acoust. Soc. Am.*, vol. 125, no. 1, pp. 372–83, Jan. 2009.

[7] S. Kerber and B. U. Seeber, "Sound Localization in Noise by Normal-Hearing Listeners and Cochlear Implant Users," *Ear Hear.*, vol. 33, no. 4, pp. 445–457, 2012.

[8] B. U. Seeber and H. Fastl, "Localization cues with bilateral cochlear implants," vol. 123, no. 2, pp. 1030–1042, 2008.

[9] J. M. Aronoff, Y. Yoon, D. J. Freed, A. J. Vermiglio, I. Pal, and S. D. Soli, "The use of interaural time and level difference cues by bilateral cochlear implant users," *J. Acoust. Soc. Am.*, vol. 127, no. 3, pp. EL87-EL92, 2010.

[10] J. W. Strutt, "On our perception of sound direction," *Philos. Mag. Ser. 6*, vol. 13, no. 74, pp. 214–232, Feb. 1907.

[11] S. A. Shamma, N. M. Shen, and P. Gopalaswamy, "Stereausis: binaural processing without neural delays.," *J. Acoust. Soc. Am.*, vol. 86, no. 3, pp. 989–1006, Sep. 1989.

[12] A. Kan and R. Y. Litovsky, "Binaural hearing with electrical stimulation," *Hear. Res.*, vol. 322, pp. 127–137, Apr. 2015.

[13] R. J. M. van Hoesel, "Exploring the Benefits of Bilateral Cochlear Implants," *Audiol. Neurotol.*, vol. 9, no. 4, pp. 234–246, 2004.

[14] R. Y. Litovsky, M. J. Goupell, S. Godar, T. Grieco-Calub, G. L. Jones, S. N. Garadat, S. Agrawal, A. Kan, A. Todd, C. Hess, and S. Misurelli, "Studies on bilateral cochlear implants at the University of Wisconsin's Binaural Hearing and Speech Laboratory.," *J. Am. Acad. Audiol.*, vol. 23, no. 6, pp. 476–94, Jun. 2012.

[15] B. S. Wilson, "Cochlear implants: Current designs and future possibilities," *J. Rehabil. Res. Dev.*, vol. 45, no. 5, pp. 695–730, Dec. 2008.

[16] I. Hochmair, "Cochlear Implants: Facts," *Med-El*, 2013. [Online]. Available: http://www.medel.com/cochlear-implants-facts/. [Accessed: 16-May-2018].

[17] A. E. Vandali, L. A. Whitford, K. L. Plant, and G. M. Clark, "Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system.," *Ear Hear.*, vol. 21, no. 6, pp. 608–24, Dec. 2000.

[18] L. K. Holden, M. W. Skinner, T. A. Holden, and M. E. Demorest, "Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy.," *Ear Hear.*, vol. 23, no. 5, pp. 463–76, Oct. 2002.

[19] P. P. Khing, B. A. Swanson, and E. Ambikairajah, "The effect of automatic gain control structure and release time on cochlear implant speech intelligibility," *PLoS One*, vol. 8, no. 11, 2013.

[20] B. S. Wilson, C. C. Finley, D. T. Lawson, R. D. Wolford, D. K. Eddington, and W. M. Rabinowitz, "Better speech recognition with cochlear implants.," *Nature*, vol. 352, no. 6332, pp. 236–8, Jul. 1991.

[21] M. Bingabr, B. Espinoza-Varas, and P. C. Loizou, "Simulating the effect of spread of excitation in cochlear implants.," *Hear Res*, vol. 241, no. 1–2, pp. 73–79, Jul. 2008.

[22] F. A. Rodriguez-Campos and M. J. Goupell, "Interaural Differences in Firing Patterns between Coding Strategies for Unsynchronized Bilateral Cochlear-Implant Freedom and N5 Processors," in *38th Midwinter Meeting of the Association for Research in Otolaryngology*, 2015.

[23] B. Laback, K. Egger, and P. Majdak, "Perception and coding of interaural time differences with bilateral cochlear implants," *Hear. Res.*, vol. 322, pp. 138–150, Apr. 2015.

[24] A. Ihlefeld, A. Kan, and R. Y. Litovsky, "Across-frequency combination of interaural time difference in bilateral cochlear implant listeners.," *Front. Syst. Neurosci.*, vol. 8, p. 22, Jan. 2014.

[25] A. Kan, H. G. Jones, and R. Y. Litovsky, "Effect of multi-electrode configuration on sensitivity to interaural timing differences in bilateral cochlear-implant users," *J. Acoust. Soc. Am.*, vol. 138, no. 6, pp. 3826–3833, Dec. 2015.

[26] B. Laback, I. Zimmermann, P. Majdak, W.-D. Baumgartner, and S.-M. Pok, "Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing.," *J. Acoust. Soc. Am.*, vol. 130, no. 3, pp. 1515–29, Sep. 2011.

[27] H. Hu, S. D. Ewert, D. McAlpine, and M. Dietz, "Differences in the temporal course of interaural time difference sensitivity between acoustic and electric hearing in amplitude modulated stimuli," *J. Acoust. Soc. Am.*, vol. 141, no. 3, pp. 1862–1873, Mar. 2017.

[28] H. Møller, "Fundamentals of binaural technology," *Appl. Acoust.*, vol. 36, no. 3–4, pp. 171–218, 1992.

[29] S. Foster, "Impulse response measurement using Golay codes," *Acoust. Speech, Signal Process. IEEE*, pp. 929–932, 1986.

[30] N. Epain, P. Guillon, A. Kan, R. Kosobrodov, D. Sun, C. Jin, and A. van Schaik, "Objective evaluation of a three-dimensional sound field reproduction system," in *Proceedings of the 20th International Congress on Acoustics*, 2010.

[31] E. A. Macpherson and J. C. Middlebrooks, "Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited.," *J. Acoust. Soc. Am.*, vol. 111, no. 5 Pt 1, pp. 2219–36, May 2002.

[32] P. C. Loizou, "Speech Processing in Vocoder-Centric Cochlear Implants," in *Cochlear and Brainstem Implants*, vol. 64, A. R. Møller, Ed. S. Karger AG, 2006, pp. 109–143.

[33] K. Chung and N. McKibben, "Microphone directionality, pre-emphasis filter, and wind noise in cochlear implants.," *J. Am. Acad. Audiol.*, vol. 22, no. 9, pp. 586–600, Oct. 2011.