

# IMAGE-GUIDED CUSTOMIZATION OF FREQUENCY-PLACE MAPPING IN COCHLEAR IMPLANTS

Hussnain Ali<sup>1</sup>, Jack H. Noble<sup>2</sup>, René H. Gifford<sup>3</sup>, Robert F. Labadie<sup>4</sup>, Benoit M. Dawant<sup>2</sup>,  
John H.L. Hansen<sup>1,5</sup>, Emily Tobey<sup>5</sup>

<sup>1</sup>Department of Electrical Engineering, The University of Texas at Dallas, Richardson, USA

<sup>2</sup>Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, USA

<sup>3</sup>Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, USA

<sup>4</sup>Department of Otolaryngology, Vanderbilt University Medical Center, Nashville, USA

<sup>5</sup>Department of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, USA

jack.h.noble, rene.h.gifford, robert.labadie, benoit.dawant@vanderbilt.edu,  
hussnain.ali, john.hansen, etobey@utdallas.edu

## ABSTRACT

Multi-channel cochlear implants (CI) leverage frequency based cochlear tonotopic mapping to map acoustic information to the cochlear place of stimulation which is primarily determined by electrode locations. Despite the fact that electrode locations within the cochlea are unique to each patient, the acoustic frequencies assigned to the electrodes by the CI processor are determined generically, resulting in a mismatch between intended and actual pitch perception. This is known to be a limiting factor for hearing outcomes with CIs. In this study, we propose a novel, image-guided CI processor programming strategy to select more optimal, patient-customized frequency assignments. The performance of the proposed strategy was evaluated using vocoder-based simulations with ten normal hearing listeners. In our simulations, our strategy results in significantly better speech recognition scores than the standard clinical strategy.

**Index Terms**— Cochlear implants, sound coding, algorithms.

## 1. INTRODUCTION

Complex temporal-spectral patterns of neural activity occur in the peripheral auditory system [1]. In natural hearing, sound stimulates auditory nerve fibers in the cochlea to induce the sensation of hearing. Auditory nerve fibers have intrinsic “characteristic frequencies” (CFs) and are tonotopically organized in the cochlea, i.e., nerve fibers located at deeper sites along the length of the cochlea have lower CFs and thus when they are simulated, lower pitched sounds are perceived. Cochlear implants (CI) exploit this natural phenomenon by providing electrical stimulation across the length of the cochlea via an electrode array, which is *blindly threaded* into the cochlear bony labyrinth during surgery. Insertion depth of the electrode array, number of

electrodes, degree of neuronal survival, positioning and proximity of electrodes to the auditory nerve fibers largely determine which tonotopically mapped groups of nerve fibers are stimulated by each electrode. Variations in the above parameters along with other physiological and cognitive factors, (e.g., age at implantation, duration of deafness and implant use, rehabilitation, to name a few) are key factors responsible for large variations in cochlear implant outcomes.

While there have been outstanding advancements in CI signal processing, there is an open-ended research issue in that once the CI surgical procedure is performed, there is a mismatch between frequency bands and the true tonotopic locations where stimulation should occur along the basilar membrane. Deeper electrode insertions generally favor improved speech recognition in CIs [2-7]. This is due to accessibility of the apical regions of the cochlea which correspond to lower frequencies and hence theoretically more low-frequency speech information (e.g., location of F0/F1 formants) can be provided (without spectral distortion). However, deeper electrode insertion has challenges of its own, e.g., insertion trauma.

Variations in electrode insertion depth result in differences in accessible tonotopic range among implant recipients, (i.e., range of CF stimulated at the corresponding electrode locations). For example, with insertion depth of 30 mm from the round window, the most apical electrode would correspond to CF of approximately 185 Hz, while a shallower insertion of 20 mm will correspond to 1170 Hz (a clear mismatch between the intended frequency of stimulation and corresponding correct location along the basilar membrane). Despite these variations in electrode insertion depth, contemporary CI sound processors *use a standard mapping strategy for all implantees* and map the full acoustic range of speech (approximately 100 - 200 up to 8500 Hz) to the tonotopic location of electrode array and

Channel no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Lower cut-off frequency (Hz)	188	313	438	563	688	813	938	1063	1188	1313	1563	1813	2063	1313	2688	3063	3563	4063	4688	5313	6063	6938
Center Frequency (Hz)	250	375	500	625	750	875	1000	1125	1250	1438	1688	1938	2188	2500	2875	3313	3813	4375	5000	5688	6500	7438
Higher cut-off frequency (Hz)	313	438	563	688	813	938	1063	1188	1313	1563	1813	2063	1313	2688	3063	3563	4063	4688	5313	6063	6938	7938
Band-width (Hz)	125	125	125	125	125	125	125	125	125	250	250	250	250	375	375	500	500	625	625	750	875	1000

Fig. 1. Default frequency allocation table for analysis bands in ACE sound processing strategy.

*simply hope* that CI users will learn to adapt to the modified map over time. Such a mismatch between freq. analysis bands of CI sound processor and the CF of the nerve fibers that are stimulated can result in frequency-place shifting (frequency offset), frequency compression, expansion, warping, or a combination of the above. These factors deteriorate spectral characteristics of the perceived sound and hence reduce speech intelligibility. Many studies have explored the effect of these mentioned spectral distortions on speech intelligibility in normal hearing (NH) listeners and CI users [7-14]. Scientific findings suggest that peak performance is achieved when full acoustic range is mapped to the tonotopic map in matched condition (i.e., analysis bands correspond to the tonotopic map of the cochlea); however, minor mismatch does not account for the significant reduction in performance [7, 12].

Contrary to these cited studies, some research groups argue that results from acute studies underestimate the effect of learning/adaptation, and that neural plasticity of the cortex can facilitate the creation of an “adapted electric map” over time, (i.e. the listener gradually adapts to the altered pattern of stimulation) [1, 11, 13]. While the extent of brain plasticity is currently unknown, current data suggests tolerance of only few millimeters [7]. It is generally agreed that large spectral distortions caused by severe frequency-place mismatch could be one of many factors responsible for low asymptotic performance as well as longer adaptation periods among implant recipients.

The strategy we present leverages these above findings and attempts to reach a compromise between frequency-place matching, frequency compression, and truncation of low frequencies. We propose a *user-customized mapping strategy*, which relies on information provided by *image analysis algorithms that operate on CT scans of individual recipients* [15]. The details of the algorithm are provided in Sec. 2, followed by acute evaluation with 10 NH participants in Sec. 3. Conclusions are presented in Sec. 4.

## 2. METHOD

In a cochlear implant, any range of frequency can theoretically be presented to any electrode [9]. Commercial CI sound processors typically map the full acoustic range from approximately 100 – 8500 Hz to the electrode array which comprises of 12 – 22 electrodes. Figure 1 shows the default frequency mapping scheme of analysis filters in Advance Combinations Encoder (ACE) [16] sound coding strategy used in Nucleus devices manufactured by Cochlear

Ltd. Since low frequencies are more critical for speech understanding, higher numbers of channels with smaller bandwidths are used to represent low frequencies. The *same* frequency allocation table is normally used in all CI users despite variations in electrode locations.

Noble *et al.* have devised a new image processing technique which accurately locates the spatial location of electrode contacts and neural stimulation sites from pre and post implantation CT scans of recipients’ cochleae [14, 17, 18]. The final output of this approach is programming-relevant information in the form of electrode distance-vs.-frequency (DVF) curves, an example of which is shown in Figure 2. Each blue or red DVF curve in the plot corresponds to an electrode in the array. A DVF curve defines the Euclidean distance from the corresponding electrode to the closest tonotopically mapped neural stimulation sites. Distance is shown on the y-axis and the tonotopic frequency of the neural sites is varied on the x-axis. Thus, a DVF curve defines the distance from different neural sites to the corresponding electrode. These curves provide insight not only into CF of each stimulation site, but also the degree of spectral overlap caused by the neighboring electrodes. A comparison of CF determined by electrode DVF curves (Figure 2) and default frequency allocation (Figure 1) reveals a high degree of frequency mismatch. In a previous clinical study, Noble and Gifford *et al.* have shown that hearing outcomes can be improved through reduction of spectral overlap artifacts using a user-customized strategy that relies on information from DVF curves [19]. However, that strategy did not attempt to address frequency mismatch artifacts. Here, we propose a new, *user-customized frequency allocation scheme* based on

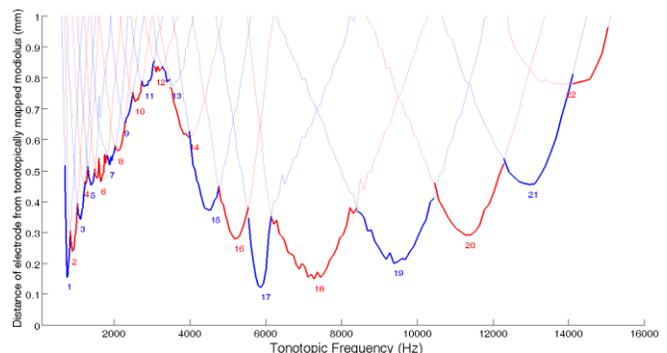


Fig. 2. Electrode distance-vs.-frequency curves of a randomly selected implant user, shown as a sequence of blue and red segments.

the DVF curve data of individual implant users that is designed to improve outcomes by reducing frequency mismatch artifacts.

The proposed frequency allocation scheme derives center frequencies of the filterbanks from DVF curves. Each curve in Figure 2 corresponds to the spatial proximity of an electrode to the nerve fibers, and minimum points on the curve represent the center CF stimulated by that electrode. We use these CFs of the stimulation sites as a reference to design analysis filter-banks. The frequency space is divided into three sub-bands:  $B_1$ ,  $B_2$ , and  $B_3$ , with frequency ranges of  $\omega_1 = [0.5\text{--}1.0]$  kHz,  $\omega_2 = [1.0\text{--}3.0]$  kHz, and  $\omega_3 = [3.0\text{--}8.0]$  kHz respectively. From DVF curves, we first determine the number of electrodes,  $n_i$ , whose CFs lie in each of the  $i=1, 2, 3$  sub-bands and then we follow the following set of procedures:

**Step#1:** If  $n_i \geq 2$ , design  $n_i$  filter-banks with center-frequencies corresponding to CFs of the electrodes (perfect matching). Similarly, design  $n_2$  filters in  $B_2$  space by perfectly matching the center frequencies of the filters with the corresponding CFs of the curves.

**Step#2:** If  $n_i < 2$ , borrow  $(2 - n_i)$  filters from  $B_2$  and map them on to the  $B_1$  space. Introduce mild frequency compression in lower-most bands of  $B_2$  (to compensate for filters allocated to  $B_1$ ) while maximizing frequency matching of the remaining filters in  $B_2$  with CFs of the curves.

**Step#3:** Design  $n_3$  filters in  $B_3$  by using logarithmic/mel filter spacing.

The aim of this 3 step rule set was to maximize frequency matching at lower frequencies (less than 3 kHz) while ensuring the lowest frequencies are not truncated. For shallow insertion depths, instead of matching frequencies and thus truncating the low frequencies, we use a mild frequency compression while maximizing frequency matching between 1-3 kHz. This is based on the speech intelligibility index (SII) which weights frequency

information in this range most critical for speech understanding [20]. In order to avoid loss of low frequencies for shallow insertions, a minimum of 2 filters are always allotted in  $B_1$  space. For deeper insertions which provide tonotopically accurate access to frequencies lower than 500Hz, filter-banks are matched according to DVF curves. Figure 3 depicts the relationship between electrode locations in the cochlea, their tonotopic frequencies, and frequency-to-place mapping in (a) the standard/default fitting technique and (b) the user-customized mapping technique proposed here. The tonotopic map is derived from DVF curves of an implant user and varies across CI recipients. Figure 3(b) clearly shows a reduction in the spectral shift and frequency-compression in the customized map as compared to the standard map; however, it is achieved at the cost of decreasing the number of analysis bands in low frequencies, which may have significant implications for speech recognition.

### 3. PROCEDURE, EVALUATION, AND RESULTS

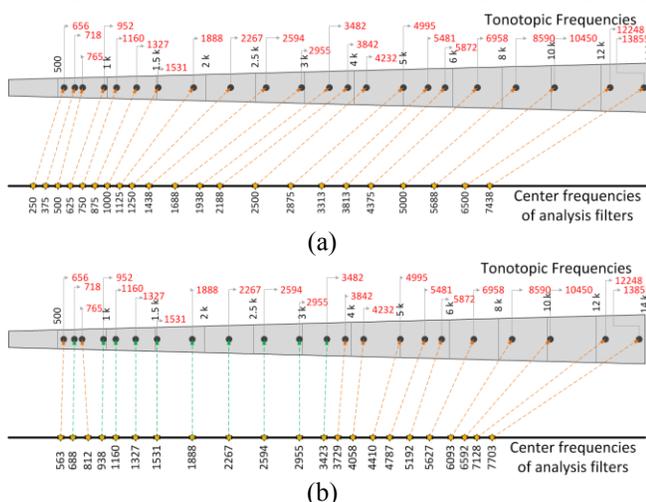
In order to evaluate performance of the proposed imaging based technique and compare it against standard methods, we followed the following experimental protocol.

#### 3.1. Experimental method

Ten normal-hearing (NH) listeners between the ages of 18-24 participated in the study. All participants were native speakers of American English, and had pure tone audiometric thresholds better than 20 dB HL at octave frequencies from 250 to 8000Hz. Each subject was tested with a single unique frequency-place map which was determined from imaging data (DVF curves) of 10 CI users (1 map/subject). CFs computed from curve minimas on electrode DVF curves were used to set center frequencies of filter banks in the following mapping conditions.

In order to simulate CI sound processing, a noise-band vocoder was implemented. The input signal was first pre-emphasized and passed through a set of bandpass analysis filters. Next, envelopes from each frequency channel were extracted via rectification and low-pass filtering. The envelope of each band was modulated with white noise and the resulting multiband signal was passed through a set of bandpass synthesis filters. The signals were finally summed up across all bands to produce a single vocoder-processed acoustic signal. Each analysis filter determines the acoustic frequency range assigned to each CI electrode, while each synthesis filter simulates the perceived sound when the corresponding CI electrode is activated and stimulates a group of auditory nerve fibers. We manipulate the characteristics of the analysis and synthesis filters to simulate the following four mapping conditions:

**Condition#1: Ideal CI position, default filter condition:** In this condition, default ACE analysis filters and identical synthesis filters are used to simulate a perfect stimulation place matching.



**Fig. 3.** An example of frequency-place mapping in (a) clinical processors, and (b) using proposed mapping strategy (figure not to scale).

**Condition#2: True CI position, default filter condition:** In this condition, we try to mimic the actual listening mechanism of CI users when using the default analysis filters. Default ACE filterbanks were used at analysis stage, whereas filterbanks derived from DVF curves were used at the synthesis stage to simulate the perceived sound.

**Condition#3: True CI position, proposed filter condition:** Custom filter-banks were designed according to each individual's DVF curve data using the methods from Sec. 2. These custom filterbanks were used as analysis filters and filterbanks derived from DVF curves were used as synthesis filters to simulate the perceived sound.

**Condition#4: True CI position, exactly matched filter condition:** Analysis and synthesis filter-banks were chosen identically from the DVF curves.

Speech recognition was assessed using four sets of test materials, namely vowels, consonants, speech in quiet, and speech at +10 dB signal-to-noise ratio (SNR) with speech-shaped noise. Vowel stimuli consisted of 12 medial vowels presented in /h/-vowel-/d/ context [21]. Consonant stimuli consisted of 20 medial consonants presented in /a/-consonant-/a/ context [22]. Recorded IEEE sentences [23] were used as the stimuli for testing speech understanding in quiet and noise. Each listener was presented 20 sentences per test condition. Each test material was presented in both male and female voices, with test material order for all test conditions randomized across subjects. The acoustic stimuli were presented in free field at 65dB sound pressure level from a single speaker in a double-wall sound booth. Performance was measured acutely without any training; however, participants were given glimpses of vocoder-processed stimuli before the start of each test material. In order to avoid any learning effects, no repetitions were allowed in any test condition.

### 3.2. Results

Figure 4 shows mean speech understanding scores for each of the four mapping conditions with different test materials. Consistent with findings from previous studies, the results here indicate peak performance with ideally matched condition (Cond#1) (i.e., full range of acoustic information is matched exactly across analysis and synthesis filter banks). However, since Cond#1 is not generally achievable in real life, the aim of the study was to compare performance of Conds #3 and #4 against Cond#2, which simulates results that are achieved with the current clinical process. Results indicate Cond#3 generally performed equal or better in all tests as compared to Conds #3 and #4, with largest improvement seen for speech in quiet (+15% improvement) followed by vowel identification (+7% improvement).

Repeated-measures analysis of variance (ANOVA) was also performed to assess the effects of mapping conditions and speech material on the speech understanding scores with an  $\alpha$  factor set to 0.05. Subjects were considered a random factor, while mapping conditions and speech material were

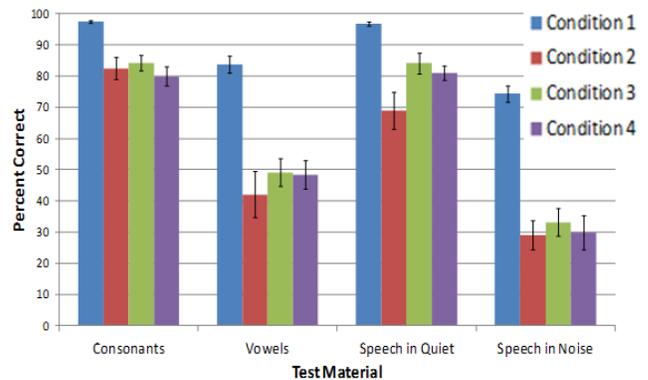
used as the main analysis factors. ANOVA revealed a significant main effect of mapping condition ( $F[3,27]=37.894$ ,  $p<0.001$ ) and test material ( $F[3,27]=97.391$ ,  $p<0.001$ ) on speech understanding scores. The interactions between mapping condition and test material were statistically significant ( $F[9,81]=12.424$ ,  $p<0.001$ ). A *Post-hoc* Bonferroni test for pair-wise comparisons between the four mapping conditions indicated statistically significant improvement with Cond#3 as compared to Cond#2 for speech understanding in quiet ( $p=0.009$ ) (no difference was observed for other 3 test materials).

## 4. CONCLUSION

While there have been outstanding advancements in cochlear implant signal processing, the lack of knowledge on spatial relationship between electrodes and stimulation targets within the cochlea has resulted in a generic fixed frequency mapping for all implantees with the hope that CI users will “learn” to interpret the incorrect frequency locations of stimulation. The proposed solution, for the first time, incorporates a CT imaging strategy to improve the CI signal processing by optimizing frequency-to-place mapping based on individual's cochlear physiology and location of electrodes. The purpose of this study is proof-of-concept validation before clinical testing with implant users. Acute results with 10 normal hearing subjects show an improvement of +15% in speech recognition indicating that user customized frequency maps can potentially aid in achieving higher asymptotic performance and possibly faster adaptation to electric hearing.

## 5. ACKNOWLEDGEMENT

This work was supported in part by grants R01DC014037 and R01DC010494 from the National Institute on Deafness and Other Communication Disorders. The content is solely the responsibility of the authors and does not necessarily represent the official views of institutes.



**Fig. 4.** Percentage correct scores for consonant and vowel recognition and speech understanding in quiet and noise with respect to mapping Conditions #1, 2, 3, 4 (from Sec. 3.1). Error bars represent standard error of mean.

## 6. REFERENCES

- [1] Q.J. Fu, and R. V. Shannon, "Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing," *The Journal of the Acoustical Society of America* vol. 105, issue 3, pp. 1889-1900, 1999.
- [2] M. F. Dorman, P. C. Loizou, and D. Rainey, "Simulating the effect of cochlear-implant electrode insertion depth on speech understanding," *The Journal of the Acoustical Society of America*, vol. 102, issue 5, pp. 2993-2996, 1997.
- [3] M. W. Skinner, *et al.*, "CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients," *Journal of the Association for Research in Otolaryngology*, vol. 3 (3), pp. 332-350, 2002.
- [4] A. Faulkner, S. Rosen, and D. Stanton, "Simulations of tonotopically mapped speech processors for cochlear implant electrodes varying in insertion depth," *The Journal of the Acoustical Society of America*, vol. 113, issue 2, pp. 1073-1080, 2003.
- [5] I. Hochmair, *et al.*, "Deep Electrode Insertion in Cochlear Implants: Apical Morphology, Electrodes and Speech Perception Results," *Acta Otolaryngol*, vol. 123 (5), pp. 612-617, 2003.
- [6] Q.J. Fu, and R. V. Shannon, "Effects of electrode location and spacing on phoneme recognition with the Nucleus-22 cochlear implant," *Ear and hearing*, vol. 20, issue 4, pp. 321, 1991.
- [7] D. Başkent, and R. V. Shannon, "Interactions between cochlear implant electrode insertion depth and frequency-place mapping," *The Journal of the Acoustical Society of America*, vol. 117, issue 3, pp. 1405-1416, 2005.
- [8] D. Başkent, and R. V. Shannon, "Speech recognition under conditions of frequency-place compression and expansion," *The Journal of the Acoustical Society of America*, vol. 113, issue 4, pp. 2064-2076, 2003.
- [9] D. Başkent, and R. V. Shannon, "Frequency-place compression and expansion in cochlear implant listeners," *The Journal of the Acoustical Society of America*, vol. 116, issue 5, pp. 3130-3140, 2004.
- [10] D. Başkent, and R. V. Shannon, "Combined effects of frequency compression-expansion and shift on speech recognition," *Ear and hearing*, vol. 28, issue 3, pp. 277-289, 2007.
- [11] Q. J. Fu, R. V. Shannon, and J. J. Galvin III, "Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant," *The Journal of the Acoustical Society of America*, vol. 112, issue 4, pp. 1664-1674, 2002.
- [11] M. J. Goupell, *et al.* "Effects of upper-frequency boundary and spectral warping on speech intelligibility in electrical stimulation," *The Journal of the Acoustical Society of America*, vol. 123, issue 4, pp. 2295 - 2309, 2008.
- [12] R. V. Shannon, F.G. Zeng, and J. Wygonski, "Speech recognition with altered spectral distribution of envelope cues," *The Journal of the Acoustical Society of America*, vol. 104, issue 4, pp. 2467-2476, 1998.
- [13] S. Rosen, A. Faulkner, and L. Wilkinson, "Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants," *The Journal of the Acoustical Society of America*, vol. 106, issue 6, pp. 3629-3636, 1999.
- [14] L. A. Reiss, M. W. Lowder, S. A. Karsten, C. W. Turner, B. J. Gantz, "Effects of extreme tonotopic mismatches between bilateral cochlear implants on electric pitch perception: a case study," *Ear Hear.*, vo. 32 (4), pp. 536-40, 2011.
- [15] J. H. Noble, *et al.*, "Image-guidance enables new methods for customizing cochlear implant stimulation strategies," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, issue 5, pp. 820 - 829, 2013.
- [16] 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 and Hearing*, vol. 21 (6), pp. 608-624, 2000.
- [17] J. H. Noble, R. F. Labadie, O. Majdani and B. M. Dawant, "Automatic segmentation of intra-cochlear anatomy in conventional CT," *IEEE Trans. Biomedical.Eng.*, vol. 58, no. 9, pp.2625 -2632 2011.
- [18] J. H. Noble, R. H. Gifford, R. F. Labadie and B. M. Dawant, "Statistical shape model segmentation and frequency mapping of cochlear implant stimulation targets in CT," in MICCAI2012, N. Ayache, Ed. *et al.* New York: Springer 2012, vol. 7511, pp.421 - 428.
- [19] J. H. Noble, R. H. Gifford, A. J. Hedley-Williams, B. M. Dawant, and R. F. Labadie, "Clinical evaluation of an image-guided cochlear implant programming strategy," *Audiology & Neurotology*, 2013 (*in press*).
- [20] American National Standards Institute (ANSI), American National Standard, "Methods for Calculation of the Speech Intelligibility Index," *Acoustical Society of America*, 1997.
- [21] J. Hillenbrand, *et al.*, "Acoustic characteristics of American English vowels," *The Journal of the Acoustical society of America*, vol. 97, issue 5, pp. 3099-3111, 1995.
- [22] R. V. Shannon, *et al.*, "Consonant recordings for speech testing." *The Journal of the Acoustical Society of America*, vol. 106, issue 6, pp. L71-L74, 1999.
- [23] IEEE Subcommittee, "IEEE Recommended Practice for Speech Quality Measurements," *IEEE Trans. Audio and Electroacoustics*, vol. AU-17, no. 3, pp. 225-246, 1969.