National Institute on Deafness and
Other Communication Disorders (NIDCD)

NIDCD

National Institute on Deafness and Other Communication Disorders

Project Title: User customization and user optimization of cochlear implants

Grant Number: R01DC010494

Final RPPR

This document has been developed to provide Principal Investigators (PIs), co-PIs, and research organizations with:

- a listing of the questions that will be asked in the new NIH project reporting format;
- assistance in planning for the submission of the report; and
- a tool to help PIs collaborate with other contributors in answering these questions, if needed.

The project reporting service on Research.gov and the associated help documentation provides more detailed instructions and contextual assistance.

Note: NIH project reports are not cumulative and should always be prepared for the specific project reporting period only.

Accomplishments

What are the major goals of the project?

Despite the many advances made in the design of novel speech coding algorithms and electrode arrays, the cochlear implant (CI) fitting process lags behind. The current fitting approach falls along the lines of "one-size-fits-all," in that it is assumed the initial fitting in the clinic should work equally well in all listening environments. Other than the psychophysical mapping (e.g., T and M levels), the speech coding algorithms and associated parameters (e.g., compression function, filter spacing) are fixed and are not optimized in any way either the patient or clinician for different listening situations (e.g., quiet, music, noise). In situations where patients wear a hearing aid in one ear and an implant in the other (bimodal users) or wear two implants (bilateral users,) the fitting is done separately for each implant, rather than jointly. The above issues limit severely our ability to tap into the full potential of existing CI devices which are not currently customized to individual users for different listening environments. This project aimed at addressing these limitations with the use of a portable research speech processor that allows users/researchers to: a) change MAP parameters, b) customize the MAP to different listening situations for maximum benefit and c) collect feedback and recordings of real acoustic signals for further analyses.

Unilateral, bilateral and bimodal users participated in speech assessment experiments to assess the true benefit of speech processing strategies that are easily customized by the users according to the environment/noise type and to fit their needs and subjective quality judgments. The overarching hypotheses is maximum benefit, measured in terms of subjective quality and/or intelligibility is obtained when the implants are customized by users to especially work in difficult listening situations experienced by implant users.

Specific project goals:

(i) User-customization and user-optimization of unilateral cochlear implants.
(ii) User-customization and user-optimization of bilateral cochlear implants.
(iii) User-customization and user-optimization of bimodal devices.

What was accomplished under these goals (you must provide information for at least one of the 4 categories below)?

[Continuous development of the research platform for cochlear implants]

In order to improve the performance and flexibility of the research platform as well to address growing research demands of the research community, a consistent effort was devoted to improving the system design, hardware, firmware, and software tools of the research platform. At the beginning of this project, Personal Digital Assistant (PDA) devices were considered state of the art in mobile computing. However, with the advent of smartphones/tablets and flexibility offered by new mobile operating systems, such as Android; it was imperative to evolve the research platform with the pace of technology. Figures below demonstrate the evolution of the research platform over the years during this grant cycle.

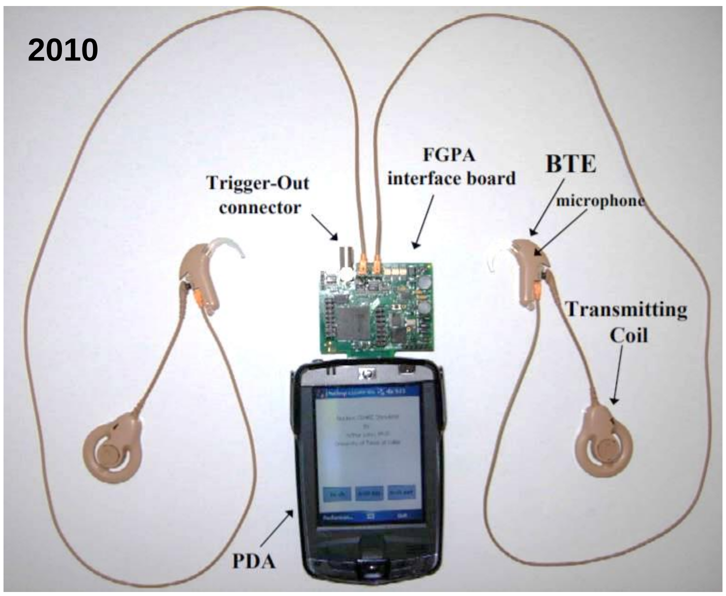


Figure 1: First Prototype of research platform with an HP PDA (2010)

Figure 2: First generation of CiPDA platform (CiPDA_v1) – (2011)

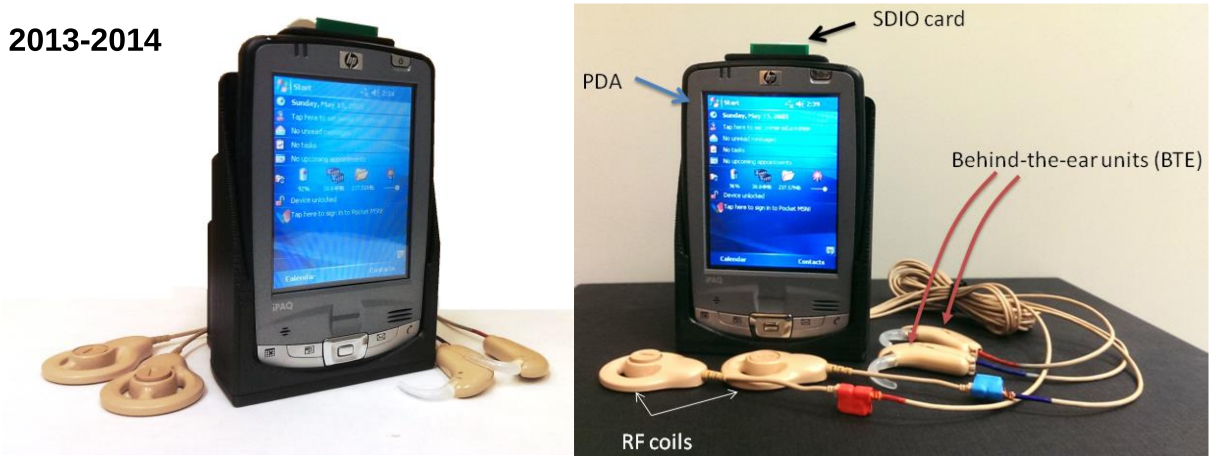


Figure 3: ciPDA version 2 (vers2) with 3D enclosure and enhanced performance (2013)

Figure 4: CCi-MOBILE-v1 – Android based re-engineered system (2015/16)

The project started with PDA devices interfaced with a Serial Digital Input/Output (SDIO) interface board to communicate with commercial implants. PDAs with Windows Mobile CE operating systems allowed sound processing algorithms to be implemented in C/C++ (fixed point environment) programming language. With system clock at 650MHz, PDAs were able to provide decent computation power for custom sound processing algorithms. Several design revisions were made throughout the project cycle to upgrade the PDA-based platform. Most notable of these design revisions include:

(i) **Updated design of SDIO interface board:** This was primarily done to improve signal quality of the acquired acoustic signal from the microphone. In the prior board, cross talk between the digital and analog circuitry was observed, which resulted in lower signal-to-noise ratio (SNR) in the acoustic signal. This impacted the overall signal processing pipeline as well as perceptual performance with the system. System upgrade was carried out to isolate digital and analog lines by providing independent power sources to digital circuitry on board (e.g., to FPGA, A/D, preamplifiers, etc.) and analog circuitry (e.g., to drive radio-frequency (RF) coils). This improvement in the board design significantly improved the overall SNR of the acoustic signals and immensely helped in running effective human experiments with the system.

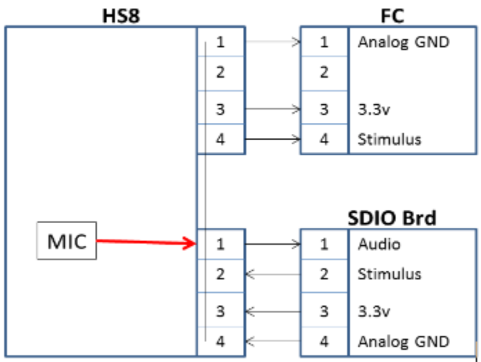


Figure 5: Original configuration of modified HS8 *BTE with common voltage source and grounds. The audio signal and stimulus signal share a common GND (e.g., ground plane) which is also used by the SDIO board ADC circuit. There was no isolation between the digital and analog circuitry, which resulted in switching noise from the digital circuitry to migrate into the analog domain through the common ground plane. Audio signal and stimulus signals also share a common 3.3v source.*

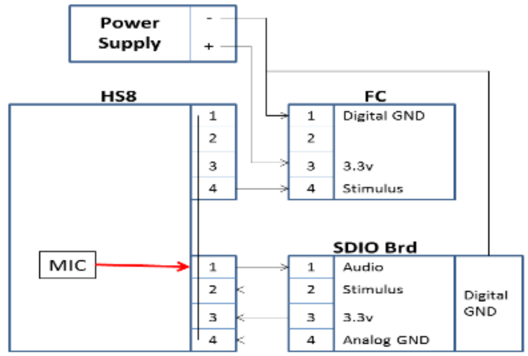
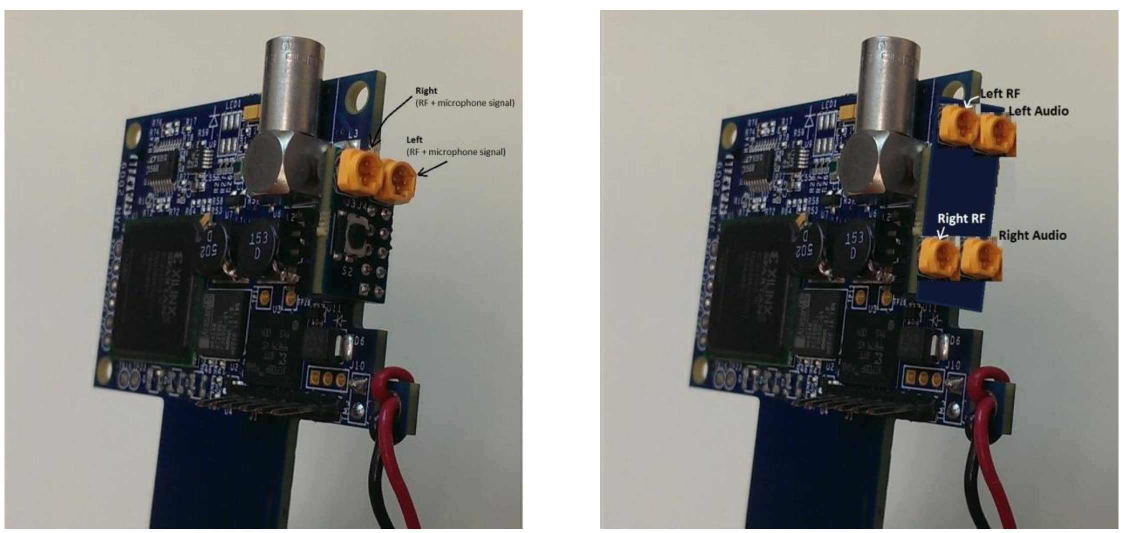


Figure 6: Updated configuration of the CIPDA setup. *BTE (microphone) is powered up by the power from the SDIO board while a separate (isolated) voltage source is used to power up the Freedom coil (FC).*



(a) Original setup of the SDIO + daughter board. Each port routes RF and acoustic signals along with power and ground.

(b) Design updates with separate ports for acoustic (audio) and RF signals for each left and right side.

Figure 7: SDIO board with daughter board ***(a) original board setup, (b) updated setup with 4 ports (2 for audio and 2 for RF).***

The platform uses HS8 BTE from Cochlear Corporation. The HS8 BTE was primarily designed for body-worn (SPRINT) processors. The inter-connects on the flex circuitry inside the HS8 BTE were modified by the team to enable it to work with an active coil. In this modified configuration, the audio signal (from the microphone) and the stimulus signal (to the Freedom coil (FC)) share a common 3.3 V source and a common ground as shown in Figure 5 below. RF interference, poor shielding, and common voltage/ground sources between the analog and digital elements on the board introduced noise in the acoustic signal. (High frequency switching in the RF, introduced cross-talk in the acoustic signal). In order to overcome this problem, the upgraded system employed a secondary power source to power up the coil independently and keeps the HS8 BTE in its original form. The block diagram for the updated setup is shown in Figure 6 below. This solution required minor changes in the design of the daughter board (Figure 7) such that the acoustic and RF traces are completely independent of each other.

The new design of the daughter board has 4 cochlear connectors, 2 for each left and right. Transmission of the RF and acoustic signals is carried out by two separate cables for each left/right. See Figure 7 for details. The new board also eliminated the need of a reset switch.

(ii) **3D enclosures:** Custom built 3D enclosures were developed to fit the system into one unit (see Fig. 3). These enclosures were necessary to run field experiments in real-world environments and immensely helped in packing up several components of the system (PDA, interface board, battery pack, cables) into a single unit.

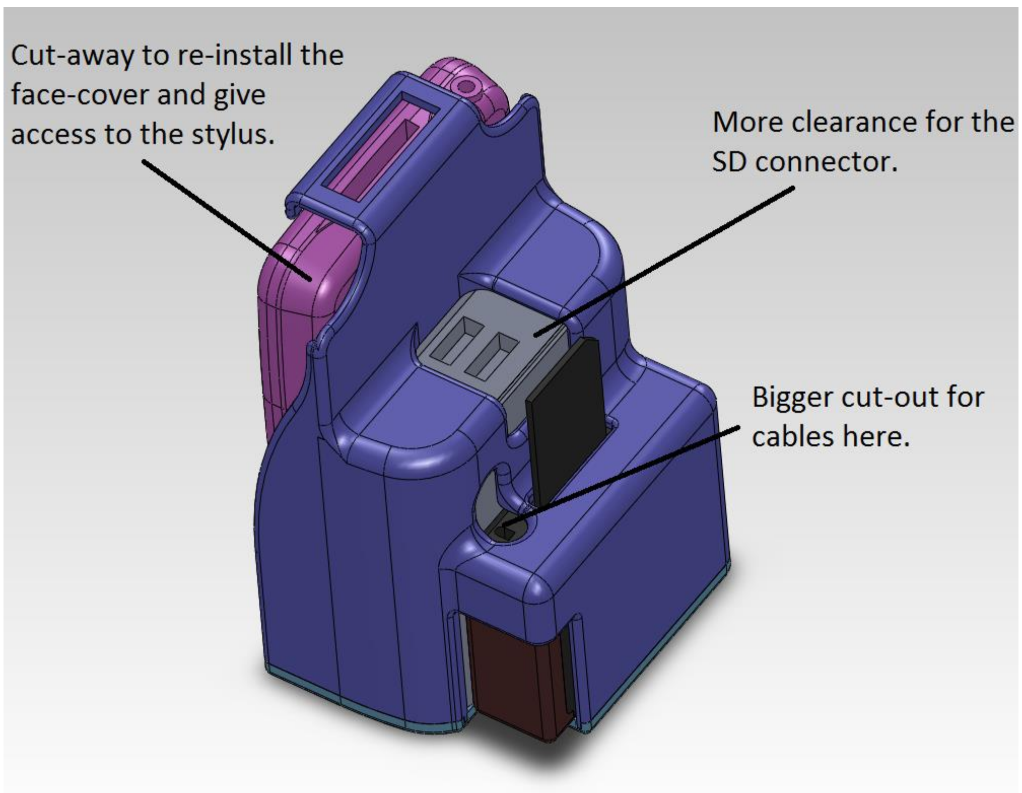


Figure 8: CAD drawing of the 3D enclosure for PDA-based research platform. The enclosure housed a PDA, a battery pack, and an SDIO interface board.

(iii) **Firmware:** Over the course of this project, several firmware updates were released to address variety of features. These updates very made in the Verilog code which resides on the FPGA and implements data transmission and communication protocols with the implant system. The firmware is locked and is not accessible to the end researchers for safety purposes. Some examples of firmware updates include:

a. Corrected Pulse Timing:
The timing of the pulses in the earlier implementations was different from that of a clinical processor. See the pulse diagram below for 5 electrodes (each filled square is a pulse and x-axis is time) in Figure 9.

In the original design of the PDA platform approach A was implemented, but clinical processors generally use approach B. That is, pulses should be evenly distributed within a time frame, with no need for null frames.

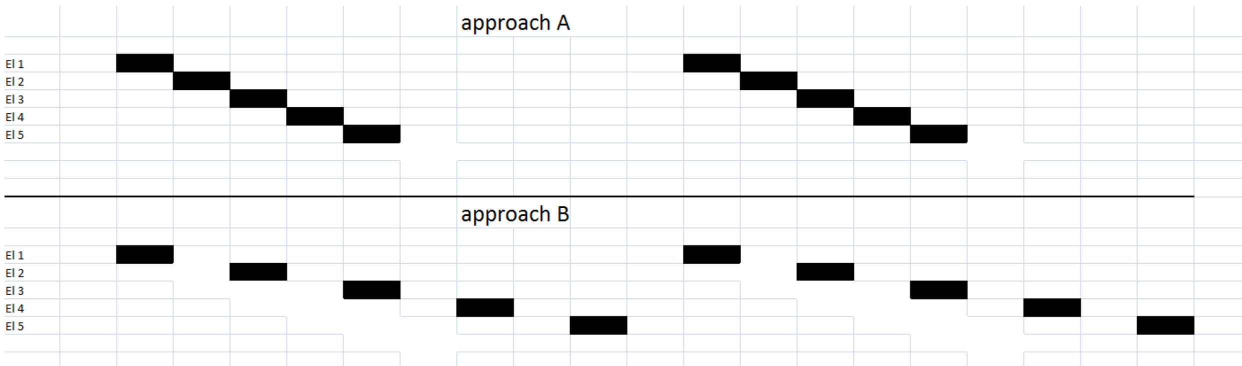


***Figure 9: Pulse Timing* - 'Approach A' is the old implementation. Approach B is that of the updated version similar to that of a clinical processor. Pulses are evenly distributed within a time frame.**

The FPGA firmware and PDA C codes were updated to change the pulse timing from approach A to approach B. The usage of null frames was stopped and instead exact timing between the pulses (*Inter-pulse duration* (**IPD**)) was employed in the update version in the FPGA to stimulate the pulses at their correct times. IPD was computed in the C code for user-specified combination of (i) stimulation rate, (ii) pulse width and (iii) number of active electrodes for each frame. IPD information is sent to the FPGA in the parameters buffer (buffer containing stimulation parameters such as implant type, stimulation protocol, pulse-width, and etcetera). The FPGA reads this information and sets the timing of the pulses based on the IPD. The state machine in the FPGA controls the timing of the pulses very precisely and only starts the stimulation at the IPD specified time.

b. Reliability of Data: Frame overlap/missing frames:

One of the biggest challenges with the design of the platform is the ability to maintain a real-time operation for sound processing needs. The platform operates on frame by frame data transfer. Each processing cycle involves, (i) capturing an 8ms data frame from A/D, ii) transmission of data packet to the processing module (PDA), (iii) sound processing for cochlear implants (e.g., CIS/ACE) in PDA, (iv) data transmission back to the interface board, (v) encoding the received data using the RF communication protocols, and finally (vi) streaming the RF encoded data to the coil. This six step process is carried out in real-time repeatedly on 8ms data packets. In order to maintain integrity of real-time processing/streaming, a number of methods were employed simultaneously. These include:

(a) High speed data transmission to and from the interface board,
(b) Optimized sound processing libraries, best programming practices for real-time applications, and use of applied DSP (e.g., fixed point programming) and
(c) Firmware (Verilog code running on FPGA) logic that controls data transmission and routing.

The firmware architecture together with best coding practices enabled real-time bilateral cochlear implant sound processing and streaming. It is noteworthy that maintaining real-time operations with mixed technologies (i.e., commercial PDAs, custom embedded circuit boards, and clinical cochlear implant) was a big challenge which was successfully addressed in the scope of this project.

c. Support for independent stimulation rate and pulse width for left and right ears

The old firmware only allowed a constant stimulation rate and pulse width combination for both left and right channels for bilateral stimulation. Firmware updates were made that allowed specification of independent stimulation rate and pulse width combinations for both left and right ears.

d. A/D and pre-amplifier configurations to improve SNR:

From hardware side, numerous challenges were faced during the course of the project. One of the challenges as indicated above was low SNR in the acoustic side. Fig. 10 below shows the block diagram of the acoustic front-end components. This includes embedded microphone circuitry (in the commercial BTE), pre-amplifier, anti-aliasing filter, and analog-to-digital converter (A/D). At hardware level, closely placed wire traces, length of traces, locations of ICs, and interconnect of components can cause unwanted interference in the very sensitive low-level signals. This together with power and ground planes sharing with digital and analog circuitries causes additional challenges. Adjustments in the transmission route of acoustic signal as well as parameters of the ICs were made throughout the project cycle to improve high SNR of the acoustic signal. By adding independent power sources for digital and analog circuitries, SNR was significantly improved (as noted above). In the later version (in CCi-MOBILE), a codec chip, with all analog front-end components embedded in a single IC, were used. Such revisions in hardware design improved the signal quality of audio streams for subsequent processing.

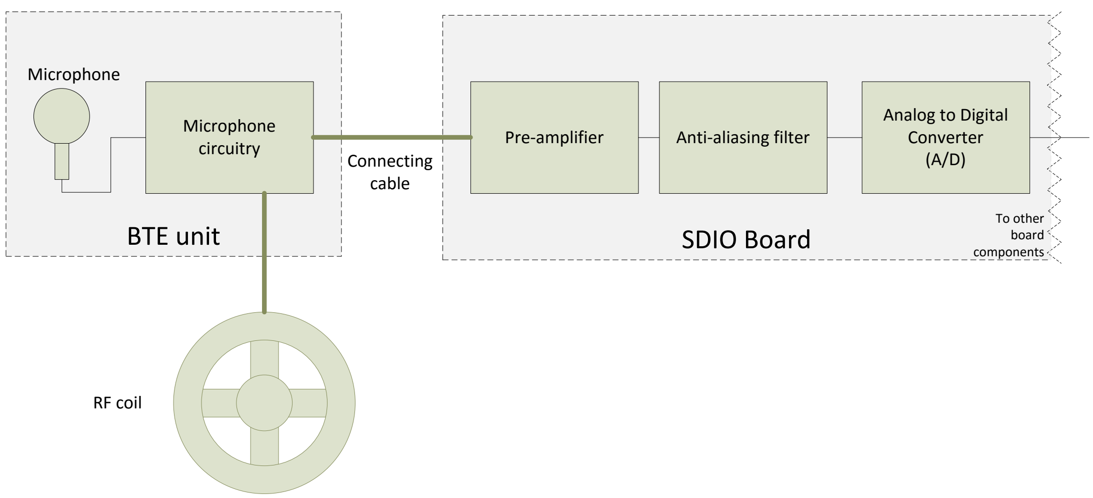


Figure 10: Block diagram of front-end microphone circuitry in the CiPDA.

e. Sampling Rate and timing of packets: 22050 kHz to 16kHz

The analog to digital (A/D) converter on the SDIO board of CiPDA setup allowed sampling at various sample rates. From the start of the project, 22050 Hz was selected as the audio sampling frequency of choice. The clinical processor, however, uses a 16000 Hz audio sampling rate. Sampling frequency of 22050 Hz was chosen to provide better signal/frequency resolution to the researchers. However, using 22050 Hz sampling rate requires adjustment in ACE coding strategy accordingly (e.g. number of samples per analysis window, FFT resolution, filter-banks, and etc.).

As per feedback from the research community and in order to make the PDA processor similar to the clinical processor, the audio sampling rate was changed to 16 kHz and the frame duration for the acoustic buffer was changed from 11.6 ms (i.e. 256 samples) to 8 ms (128 samples). This required rewriting of firmware state machines, which was a non-trivial task. The Verilog code was re-written to provide 8 ms frames (128 samples) for each left/right side at sample rate of 16 kHz. This required modifications in the signal processing strategies to address these changes accordingly. Table I lists the modifications made in the ACE signal processing chain.

Table 1: Difference in signal processing parameters for 22050 Hz and 16000 Hz sampling rates for ACE strategy.

Original Setup	Updated Setup (as used in clinical processors)
Fs = 22050	Fs = 16000
Buffer Size = 11.6 ms (256 samples)	Buffer Size = 8 ms (128 samples)
Analysis window = 8 ms (176 samples)	Analysis window = 8 ms (128 samples)
NFFT = (256 point FFT)	NFFT = (128 point FFT)
Adjusted Filter bank (bin widths, and gain)	Filter bank (bin widths, and gain) similar to clinical processor

f. System Calibration:

In order to make the output of the CiPDA processor similar to a clinical processor, extensive system calibrations were carried out. A hearing-aid analyzer with sound box (FONIX FP40 Hearing Aid Analyzer from Fry Electronics Inc.) was utilized to calibrate the system at pure tones. With our current system, we were able to achieve following characteristics:

Acoustic stimuli	Sound Pressure Level (SPL)	Electric output
1 kHz	65 dB SPL	Maximum Comfortable Level (MCL)
1 kHz	33 dB SPL	Threshold Level (THR)

Maximum Comfortable Level (MCL) is the maximum current level that is comfortable to perceive specified in a patients MAP. Threshold Level (THR) is the minimum current level that is perceived by the CI user. Both THR and MCL values are user specific.

The above parameters were provided by Cochlear Limited and are the same used in the clinical processor.

System was also analyzed at varying SPLs for:

- narrow band noises,
- pure tones at varying frequencies,
- complex signals such as chirps, and
- speech signals.

In doing this, the output current levels and the electrodogram outputs were carefully analyzed and compared with the clinical system.

A series of calibration tests were conducted on the updated system to ensure correct output. Following are details about the calibration protocols adopted.

1. Frequency Mapping

In order to ensure that correct electrodes are stimulated corresponding to the acoustic stimuli, pure tones at the center frequencies of the filter banks were presented. The outputs at each electrode were observed using an implant-in-a-box emulator together with an oscilloscope and also with the CIC4 Decoder Implant Emulator (DIET) box. The stimuli response to the tones sweeping from 200 Hz – 8000 Hz was verified.

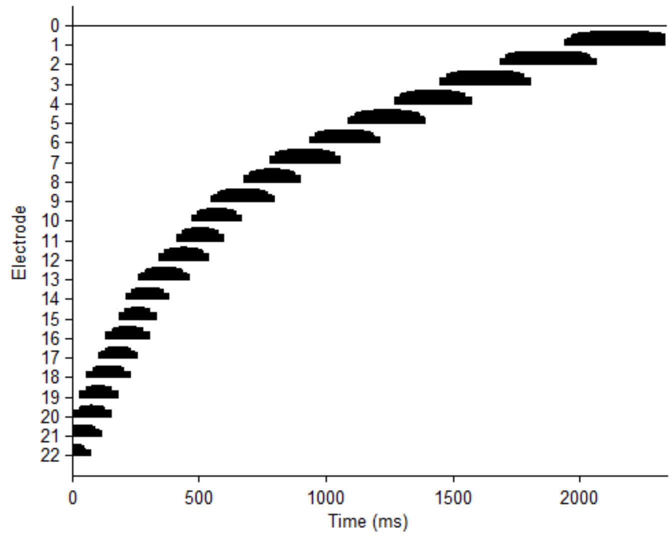


Figure 11: Response to a chirp signal with frequency sweep 200 Hz – 8 kHz

2. Loudness Mapping

The input/output characteristics (also known as default calibration values) of a clinical system are as follows:

- For a 1 kHz pure tone presented at 65 dB SPL (C-SPL), the output level at electrode 16 (the center frequency electrode) is at MCL (Maximum Comfortable Value specified in subject's MAP) value at default sensitivity level, and
- For a 1 kHz pure tone presented at 35 dB SPL (T-SPL), the output level is at THR (Threshold) level at electrode 16 at default sensitivity level.

The CIPDA v2 was calibrated to operate on the default calibration values on specifying the sensitivity values according to each individual unit. We observed small deviations in the intensities of the recorded input acoustic signals with each hardware unit. The difference primarily depends on the intensity of the acoustic signal reaching the microphone. HS8 BTE is very sensitive to small changes (distance, direction, reflections). Very subtle changes in the direction of the BTE may reflect in noticeable difference in recorded signal intensities. That is why, calibration was done on each unit and default sensitivity levels were computed for each based on the average RMS value of 1 KHz tone recorded at 65 dB SPL.

Two sliders for each left and right side were provided in the real-time application on the PDA screen, i) Sensitivity level (ranging from 0 – 20), and ii) Volume level (ranging from 0 – 10). Changing the sensitivity level amplifies or attenuates the input acoustic signal. Higher sensitivity would capture more ambient sounds. Sensitivity values greater than 10 have an effect of signal amplification, while sensitivities less than 10 would attenuate the input signal. Table 1 provides the resulting gain in the input signal corresponding to the sensitivity values.

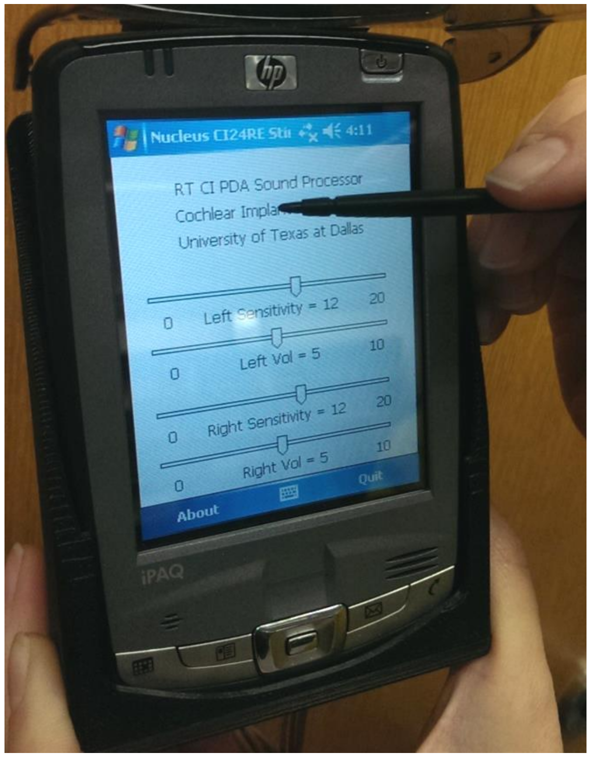


Figure 12: Sensitivity and volume sliders on the PDA screen for real-time mode.

Table 2: Effect of sensitivity values on the input gain of the acoustic signal

Sensitivity Value	Gain (dB)
0	No signal
1	-20.00 dB
2	-13.98 dB
3	-10.46 dB
4	-7.96 dB
5	-6.02 dB
6	-4.44 dB
7	-3.10 dB
8	-1.94 dB
9	-0.92 dB

Sensitivity Value	Gain (dB)
10	0 dB
11	0.83 dB
12	1.58 dB
13	2.28 dB
14	2.92 dB
15	3.52 dB
16	4.08 dB
17	4.61 dB
18	5.11 dB
19	5.58 dB
20	6.00 dB

A change in volume level would globally increase/decrease the current levels to a set percentage value of the electrical dynamic range. This feature is similar to the clinical processor. The default value is set at 5. Volume level greater than 5 would increase the loudness (current levels would not increase beyond the MCL value specified in the MAP), and volume level less than 5 would decrease the loudness value.

Loudness calibration was done in multiple ways to ensure we carry out the calibration process comprehensively: 1) Near Field calibration, 2) Far Field calibration, and 3) Test Box. Following are details on each of the test condition:

Standard (Far, free Field) calibration:

BTE was placed at a distance of 1 m from the speaker in the sound booth and a 1 kHz pure tone was presented at 65 dB SPL in the sound booth. The raw acoustic signal from the microphone was recorded, and the mean RMS value was computed. For each system, the mean RMS value was recorded over several trials. The average RMS value provided default sensitivity levels for each unit.

Near Field calibration:

For a directional microphone, such as the one in HS8 BTE, it is important to ensure that the microphone is providing directional benefit. The near-field forward response was measured by positioning the BTE at a distance of 12 inches from the speaker as shown in Fig. 8 below. The mean RMS value of the recorded signal (acoustic stimuli: 1 kHz pure tone presented at 65 dB SPL) for each system was recorded over several trials.

Figure 13: Near-field measurements at a distance of 12 inches from the speaker

Test-box calibration:

A portable hearing aid analyzer (Fonix FP40, Fig. 9) was used to make calibration measurements in the test box. The BTE was placed inside the test box and variety of signals, including pure tones, speech, and composite signals, were presented. The microphone signal was recorded, RMS values were computed and the corresponding outputs in current levels were observed and recorded at different electrodes using an implant-in-a box emulator/oscilloscope and the DIET box. The presentation level of the acoustic signal was varied from 35 dB SPL to 70 dB SPL and the corresponding change in the current levels was observed. A perfectly calibrated system saturates at 65 dB SPL (1 KHz pure tone) producing current levels at MCL value at the corresponding electrodes.

(a) *(b)*

Figure 14: (a) Fonix FP40 hearing aid analyzer used for test box calibration. (b) Test Box with BTE placed inside

g. Software:

The complete signal flow in the PDA-based speech processor is shown in Fig. 15. The acoustic signal is picked up by the microphone located in the BTE and sent to the FPGA interface board via the headset cable. The interface board samples the signal binaurally at a rate of 16 kHz/channel and sends the sampled (digital) signal to the PDA via the SD slot. The PDA processes the digital signal via a speech coding algorithm which produces stimulus data comprising of stimulation parameters and an array of current levels, stimulation mode and respective active electrodes. The stimulus data is sent to the SDIO board which transmits it to the implant using RF protocols specific to the implant.

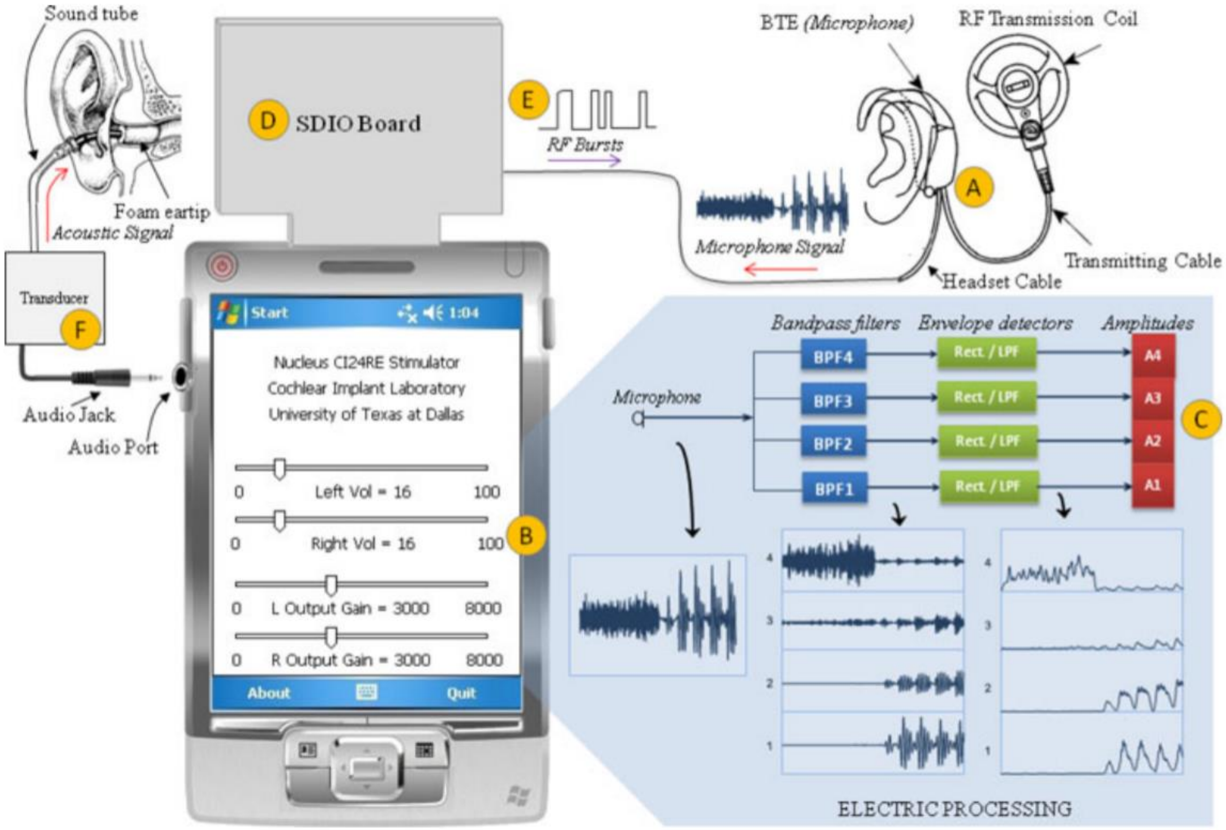


Figure 15: Signal Flow in the PDA-based speech processor.

Advanced Combination Encoder (ACE) strategy is one of the most widely used sound coding algorithm. It is used in Freedom processors developed by Cochlear Corporation. ACE strategy is implemented in fixed-point C in the PDA. Figure A1 illustrates signal flow in ACE strategy. PDA acquires 8 ms (128 samples) frames from A/D at 16 kHz per left right channel. The signal is first passed through a 1st order pre-emphasis filter. The signal is then buffered into overlapping analysis windows (each 8 ms) and 128 point Hann window is applied (Fig. A1, Step 3). The overlapping factor depends on the channel stimulation rate. In the current implementation, analysis rate is equal to the channel stimulation rate. Block Shift is defined as:

$$Block\ Shift = ceil(\frac{Audio\ Sample\ Rate}{Analysis\ Rate})$$

The Hann window $w(m)$ is defined as:

$$w(m) = 0.5\left[1.0 - \cos\left(\frac{2m\pi}{L-1}\right)\right] \qquad m = 0, \dots, L-1$$

A 128-point FFT is then performed, which yields bin centre frequencies (f_c) that are linearly spaced at multiples of 125 Hz. Because the input signal is real, the output has what is known as Hermitian symmetry, only bins 1 to 64 are used to calculate magnitude –squared spectrum as below:

$$X(k)^2 = X_R^2(k) + X_I^2(k) \qquad k = 1, \dots, L/2$$

Squared magnitude response is then appropriately weighted by a weights matrics which determines the frequency boundaries of each filter channel, according to the tables in Appendix A. The nth filter channel envelope is:

$$Y(n) = \sqrt{\sum_k a(n,k)X^2(k)} \qquad n = 1, \dots, N$$

To allow the shape of the frequency response to be adjusted, a vector of variable filter channel gains g(n) is also defined. The final filter bank output vector is:

$$Z(n) = G(n)Y(n) \qquad n = 1, \dots, N$$

Each analysis of an input block of samples produces one vector of N filter bank envelope samples. The output is sorted and the *nmaxima* outputs are selected. The selected output is compressed via logarithmic compression and finally the output is mapped to the current levels using the subject's MAP parameters (MCL and THR values).

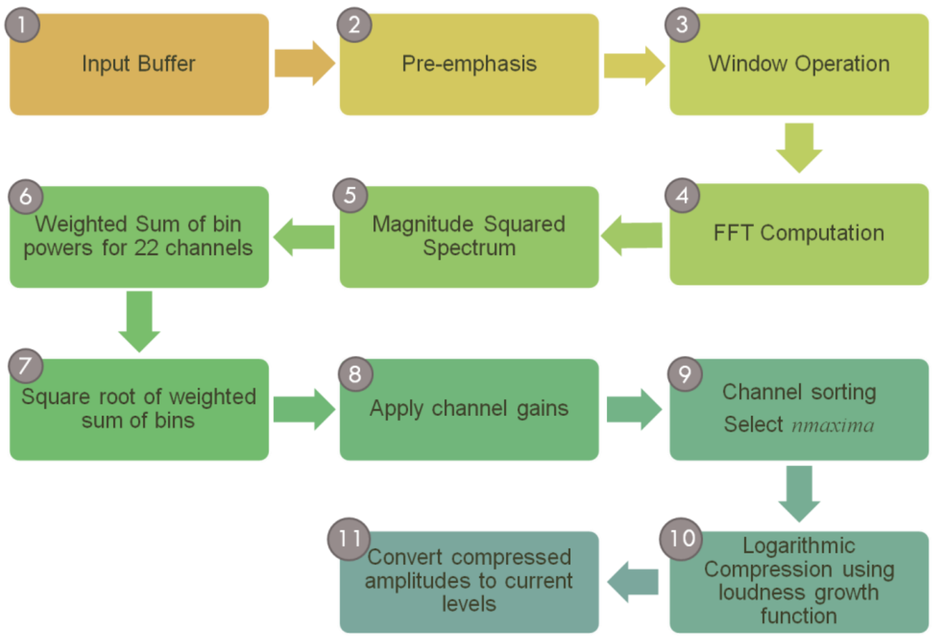


Figure 16: Sequence of steps in ACE processing strategy

Software updates were made at two fronts:

a. PDA applications: CIS and ACE sound processing algorithms were implemented in the CiPDA platform to convert acoustic signal to electric output patterns for Nucleus implant system. These sound processing strategies were programmed in a combination of C, C++ and assembly language to maintain real-time operation. In this regard, optimized primitive libraries from Intel (Intel IPP libraries for optimized signal processing) for ARM-based processors were utilized to perform high-computation tasks. In addition, fixed-point programming was adopted to cut down the computation cost. Other best practices in algorithm design and coding were utilized to improve real-time performance of the system. Table 2 below shows sequence of steps and time taken to perform computations with HP PDA (625MHz). Each processing cycle required 1.2ms for a 128 sample window.

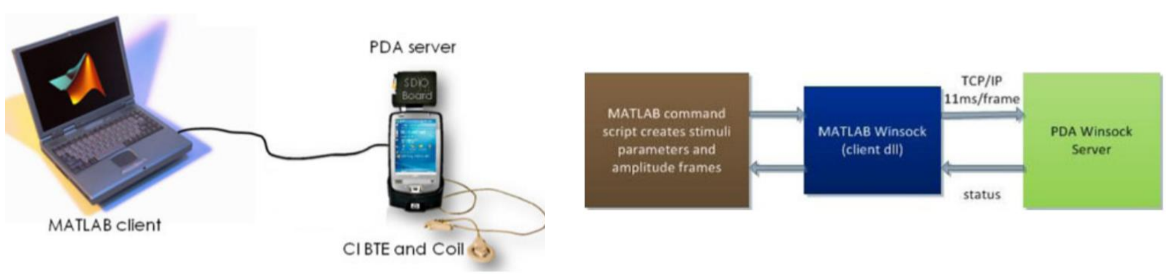
Table 3: Sequence of processes for ACE algorithm along with the PDA processor's time profiling for each step.

Process	Time (μs) per analysis window[1]	Time (μs) per frame[2]
1. Apply a Blackman window.	4.43	124.1
2. Perform a 128-point FFT of the windowed sub-frame using the IPP FFT routine [36].	9.09	254.5
3. Compute the square of the FFT magnitudes.	0.67	18.7
4. Compute the weighted sum of bin powers for 22 channels. The resulting output vector contains the signal power for 22 channels.	15.36	430.1
5. Compute the square root of weighted sum of bins.	2.63	73.6
6. Sort the 22 channel amplitudes obtained in step (5) using the shell sorting algorithm. Select the *n* (of 22) maximum channel amplitudes, and scale appropriately the selected amplitudes.	9.41	263.4
7. Compress the *n* selected amplitudes using a loudness growth function given by: $y = \log(1+bx)/\log(1+b)$, where x denotes the input signal from step 6 and b is a constant dependent on clinical parameters: base level, saturation level and Q value.	0.86	24.0
8. Convert the compressed amplitudes, y (step 7) to current levels: $I = (C-T)y+T$, where I denotes the current level, T denotes the threshold level and C denotes the most-comfortable loudness level.	0.82	22.9
Total	43.26	1.21×10^3

[1]Analysis window: 128-sample (5.8 ms) analysis window at the sampling rate of 22050 Hz.
[2]Frame: Signal is initially buffered in 255-sample (11.6 ms) frames. Subsequent processing is done on 128-sample analysis windows. Profiling statistics reported in this table were carried out at a fixed stimulation rate of 2400 pulses per second per channel, with a block shift of 9 samples on each frame. The total number of analysis windows processed per frame are 28. Steps 1 – 8 are carried out for each of 28 analysis windows per frame.

In addition to maintain real-time operation, equally important is the reliability of the output. Implementations of the strategy were compared with clinical implementations provided by Cochlear Corp. to ensure correct stimulus patterns were achieved with a given set of input. In this regard, calibration process, as detailed above, together with implant-in-a-box testing played a critical role.

b. MATLAB applications for benchtop/offline processing: The offline version of the PDA processor was based on a PC running MATLAB, where all processing took place while the PDA acted as an interface to the implant. The software architecture was designed such that the PDA acted as a server which accepts the incoming connections and the PC acted as a client with MATLAB as a front-end as shown in Fig. 17. Therefore, the overall design can be broken down into three main software components. (i) Server running on the PDA. (ii) MATLAB client (.mexw32 or .mexw64 dll) called from the MATLAB front-end. (iii) MATLAB front-end running on the PC. The server client interface was based on Windows Sockets (Winsock) API which is a technical specification that defines how Windows network software should access network services, especially TCP/IP. Fig. 17b shows the transfer of parameters and amplitudes from MATLAB to the PDA and the status returned from the PDA to MATLAB.



(a) *(b)*

Figure 17: (a) High-level diagram of PDA in offline mode. (b) Device connectivity and data exchange in offline mode.

Flexibility to program in MATLAB in offline mode allowed us (and researchers) to design applications targeted to various behavioral, psychoacoustical, and psychophysical experiments which were not easily possible with conventional research interfaces. For this purpose a software package was developed which included a versatile suite of applications to conduct a variety of experiments. Some examples are scripts for sound coding algorithms (e.g., CIS, ACE etc.), applications to conduct listening experiments with CI users, experiments for modulation detection, electrode pitch matching, finding interaural time differences (ITDs), and interaural intensity level differences (ILDs). This framework of routines, scripts, and applications was developed as a reference for the researchers to develop custom experiments according to their own research needs. The platform was also successfully integrated with Percept, which is a software package developed by Sensimetrics Corporation to facilitate the design and assessment of sound processing strategies. Percept offers a wide range of psychophysics experiments which can easily be performed on human subjects in the lab environment.

(iv) MIGRATION TO ANDROID PLATFORM:

2014 onwards, efforts were devoted to upgrade the device to newer generation of smartphones, specifically Android devices. This was done to leverage growing computing capabilities of newer generation of smartphones and because PDAs were getting obsolete. It also helped in migrating from an SDIO-based connection interface to a standard USB-based connection. The newly designed platform was called Costakis Cochlear Implant Mobile (CCi-MOBILE) research interface. Most notable of improvements include:

(a) Re-designed interface board: The new interface board has a high quality 4-channel A/D codec, Xilinx Spartan 6 FPGA, USB interface, a fully integrated Wi-Fi transceiver for wireless data transmission.

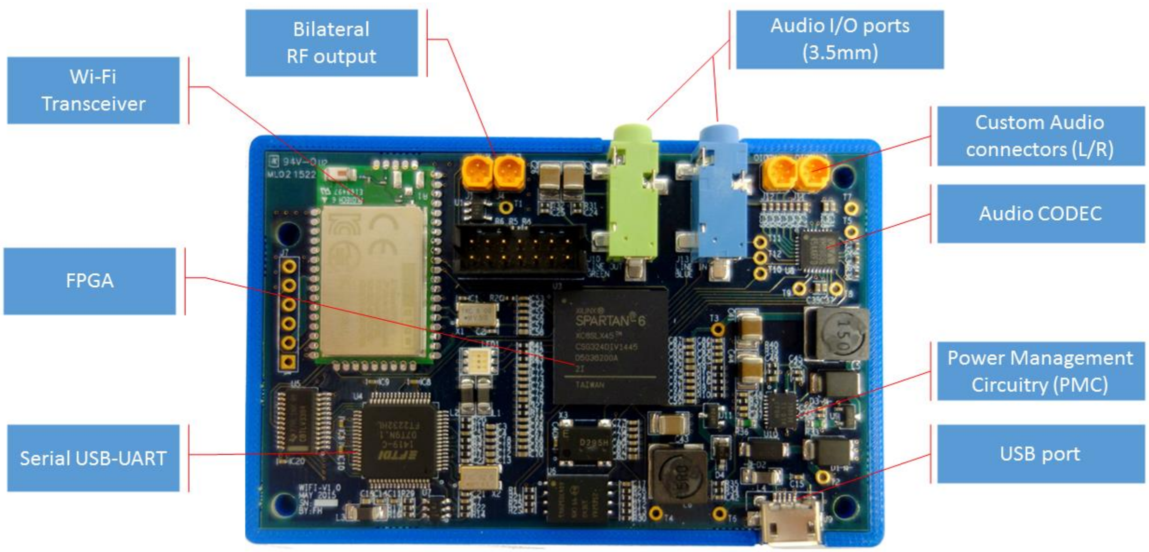


Figure 18: CCiMOBILEv1 interface board

FPGA: The central processor of the interface board is a Field Programmable Gate Array (FPGA) from Xilinx (XC6SLX45). FPGA controls the data flow in the system. It is responsible for synchronized capture of audio frames from the codec, two-way data transfer, encoding of RF data, and synchronized delivery of stimulation pulses to the RF coils. The Wi-Fi and USB transmission blocks communicate with the core FPGA using UART standard protocol, while the audio codec uses the SPI protocol for data exchange. In addition to the RF encoding, FPGA firmware also realizes safety checks for safe stimulation delivery. The FPGA is programmed in Verilog, locked, and not accessible to the researchers.

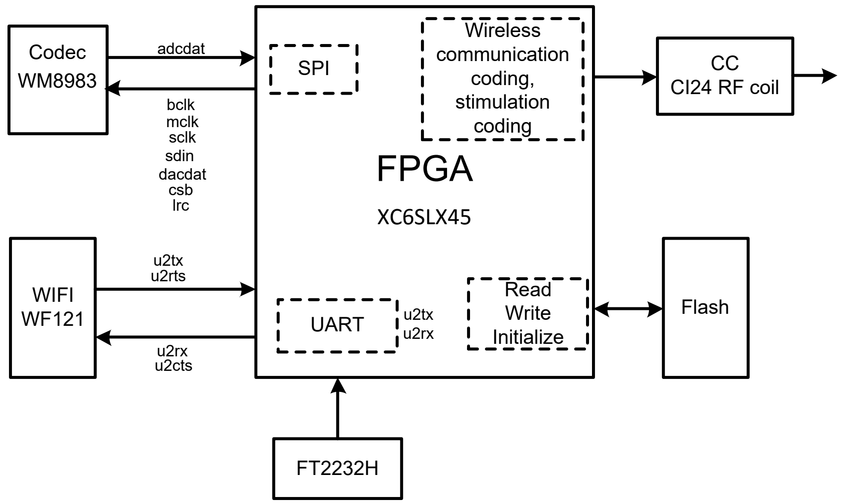


Figure 19: System level block diagram of FPGA and interfaced components showing data exchange between components

Audio Codec: The audio codec from Wolfson Microelectronics (WM8983) is a highly integrated input/output device designed for mobile computing and communications. The device integrates preamps for stereo differential mics, and includes drivers for speaker, headphone and differential or stereo line output. External component requirements are reduced as no separate microphone or headphone amplifiers are required. Advanced on-chip digital signal processing includes a 5-band equalizer, a mixed signal Automatic Level Control for the microphone or line input through the ADC as well as a purely digital limiter function for record or playback. A programmable high pass filter in the ADC path is provided for wind noise reduction and an IIR with programmable coefficients can be used as a notch filter to suppress fixed-frequency noise. Key features of the audio codec are as follows:

ADC	**4 channels**
SNR	87 dB
AGC	Yes
Interfaces	PCM, I²S
Power Supply	1.8 – 3.6 V
Package	7 × 7 mm 48-pin QFN

Wi-Fi Transceiver: The Wi-Fi transceiver from Bluegiga (WF121) is a stand-alone Wi-Fi module that provides fully integrated 2.4GHz 802.11 b/g/n radio, TCP/IP stack and a 32-bit micro controller (MCU) platform for simple, low-cost and low-power wireless IP connectivity. It also provides flexible peripheral interfaces such as SPI, I2C, ADC, GPIO, Bluetooth co-existence, and timers to connect various peripheral interfaces directly to the WF121 Wi-Fi module.

Key Features:

- 2.4GHz band IEEE 802.11 b/g/n radio
- Radio performance: TX power: +17 dBm, RX sensitivity: -97 dBm
- Host interfaces: 20Mbps UART, USB on-the-go
- Peripheral interfaces: GPIO, AIO and timers, I2C, SPI and UART, Ethernet
- Embedded TCP/IP and 802.11 MAC stacks: IP, TCP, UDP, DHCP and DNS protocols, BGAPI host protocol for modem like usage, BGScriptTM scripting language or native C-development for self-contained applications
- 32-bit embedded microcontroller: 80MHz, 128kB RAM and 512kB Flash, MIPS architecture
- Fully CE, FCC and IC qualified

USB Interface: The FT2232H is a USB 2.0 Hi-Speed (480Mb/s) to UART/FIFO IC that is used for direct connect mode to interface PC and portable USB-based devices with the board via a micro-USB connector.

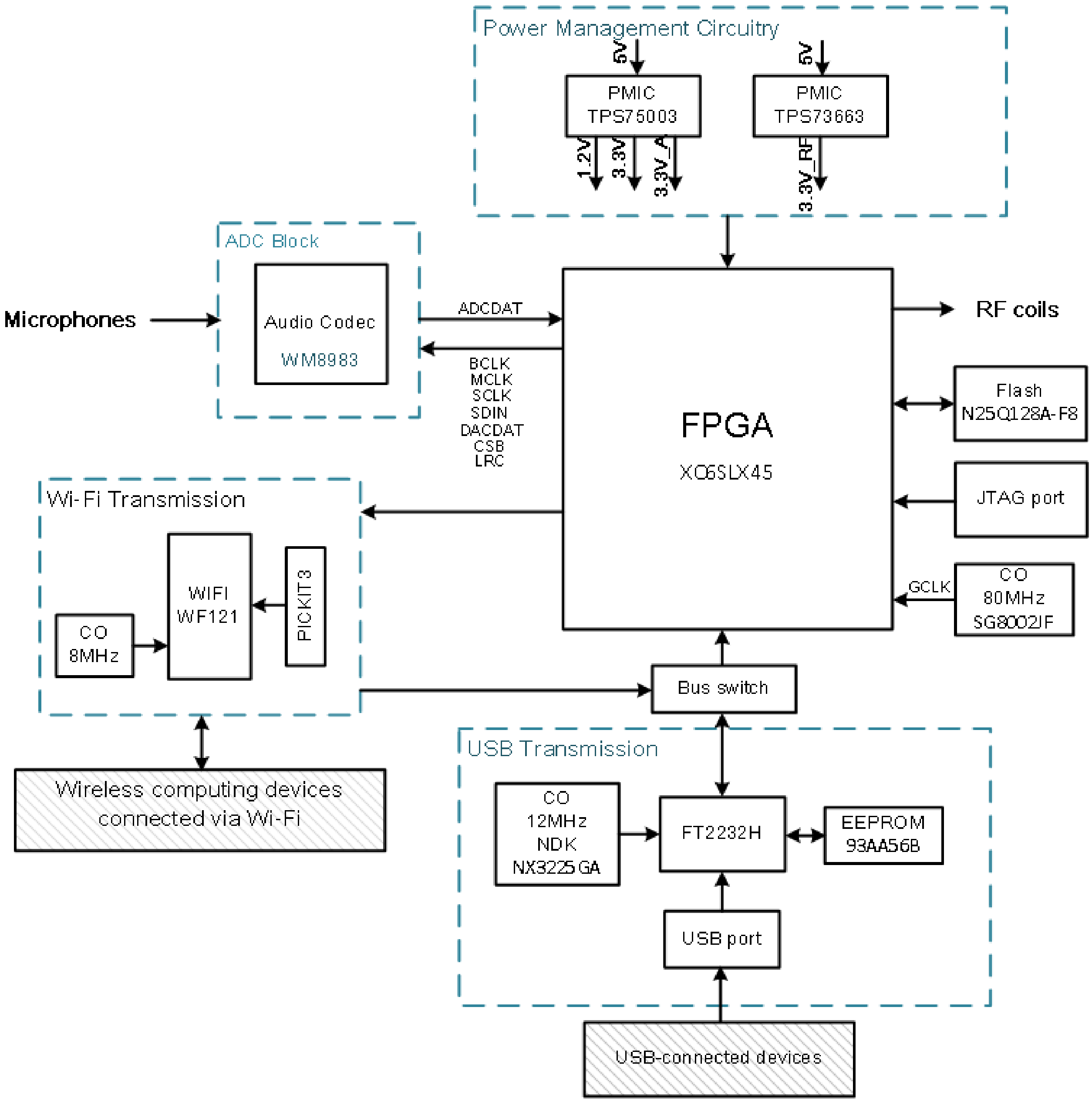


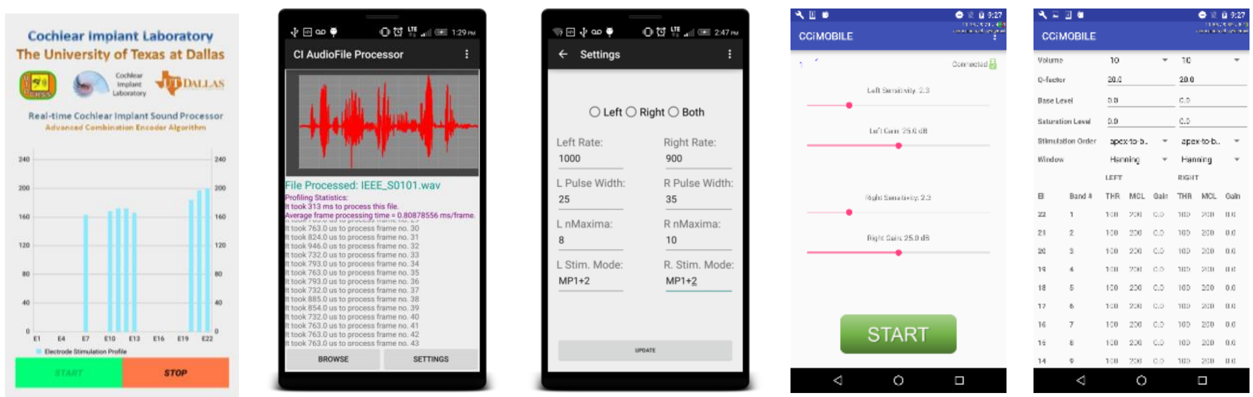
Figure 20: System block diagram of CCi-MOBILE interface board

Power Management: The circuit board runs from a 5-V battery which connects to the micro-USB port of the board. The power management circuitry comprises of two main components: TPS75003 and TPS73663. The TPS75003 is an integrated triple supply power management IC (PMIC) for the Xilinx Spartan FPGA and regulates voltage and current levels of the digital part of the circuit board. The TPS73663 is a low-dropout

(LDO) regulator and is used to provide stable power to the RF analog circuitry. The use of two separate LDOs ensures lower cross-interference between the digital and analog circuitries.

(b) Android-based mobile applications for sound processing

One of the major upgrades to the system is the use of commercially available smartphones and tablets for sound processing needs. In the upgraded platform, sound processing routines are implemented in JAVA as separate applications (Apps). Custom applications were developed for different sound processing and experimental needs. Some of these applications are shown in Fig. 13 below:



(a) (b) (c) (d) (e)

Figure 21: Android apps for different experimental needs. *(a) Real-time ACE sound processing application, (b) Audio-file (wave files) processing application. (c) Setting tab on the App. (d) User-defined inputs for volumes and sensitivity controls. (e) screen showing fully customized MAP.*

(c) PC/MATLAB-based processing capabilities:

The platform can be connected with a PC via a USB cable. In this mode, PC can act as a sound processor. PC can be configured to perform signal processing tasks either in real-time mode as well as in offline mode. MATLAB-based applications were developed to address multitude of research needs. Some examples of these applications are shown in the figures below (Figures 22 and 23).

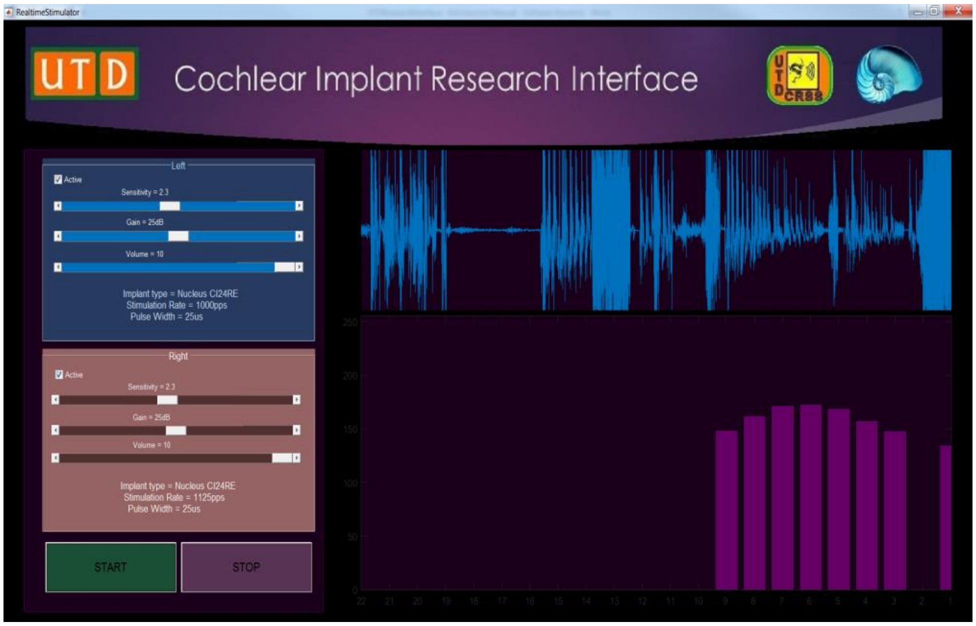


***Figure 22: Realtime ACE processing:** This application performs real-time ACE processing on the incoming acoustic signal frames.*

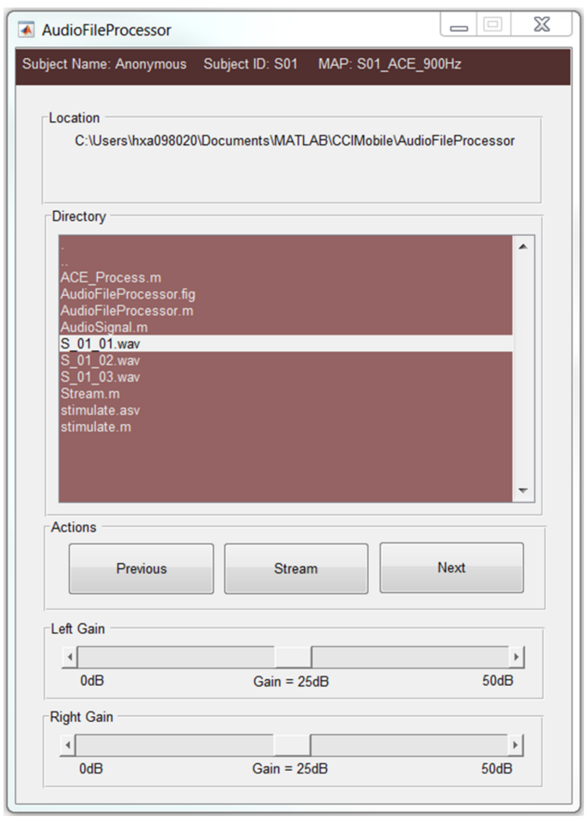


***Figure 23: Audio File Processor:** This application performs ACE processing on the audio wave files with '.wav' format.*

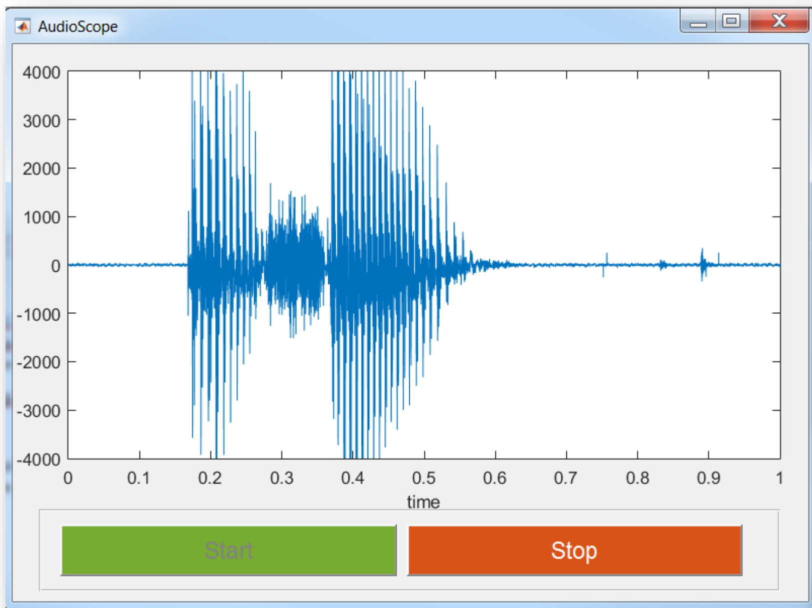


Figure 24: AudioScope is a simple application to visualize the incoming acoustic signals from the BTEs. It displays the acoustic waveforms in real-time as they are captured.

Figure 25: Audio Recorder application records the incoming acoustic signals from the BTEs and store the audio in an audio file on the PC.

[Evaluation of research platform with human subjects]

Evaluations with CiPDA platform – to verify the functionality and safety with human subjects

The PDA platform underwent rigorous testing to ensure that it met all safety criteria of a clinical processor. The output levels at the electrodes were verified extensively using 1) an implant emulator and 2) a decoder and implant emulator tool box, from Cochlear Corporation. Electrodograms produced by each processor were compared to ensure similarity for the same input.

The platform is FCC and IEC compliant. An IDE application for the PDA-based research processor was approved by the FDA in May 2011 for evaluation with human subjects. Since then, the platform has been tested with unilateral, bilateral, and bimodal CI users. The results presented below are from an acute study (i.e., users were allowed to wear the processor for a few hours in the lab environment) with ten CI users. The aim of the study was to evaluate the performance of the PDA-based speech processors on a speech intelligibly task in quiet and noisy conditions and compare the performance against the users' own clinical processor.

A total of ten CI users participated in this acute study. All participants were adults and native speakers of American English with post lingual deafness with a minimum of 1-year experience with CI(s) from Cochlear Corporation. Five subjects were bimodal CI users, i.e., wearing a cochlear implant in one ear and a hearing aid in the other.

All subjects were tested with three processor conditions: clinical, PDA-based offline (PDA_Offline) and PDA-based RT (PDA_RT) processors. The intelligibility scores from their clinical processor were used as benchmark scores. Tests with the subjects' clinical and the PDA_RT processors were conducted in free field in a sound booth at an average presentation level of 65 dB SPL. Speech stimuli for the PDA_Offline were presented via audio files on the PC. In all cases, sensitivity level and volume adjustments were completed on the respective processors. The subjects' everyday MAP was used for all the test conditions. Environmental settings (such as autosensitivity, ADRO, and BEAM), if active, were not disabled on the subjects' clinical processors. These settings were not implemented in the PDA-based processors. A short training with the PDA-based processor was carried out before each test. In addition to electric alone, bimodal subjects were tested for acoustic alone and EAS with both types of processors. In this study, no audio processing for acoustic stimulation was done.

The speech stimuli used for testing were recorded sentences from the Institute of Electrical and Electronics Engineers (IEEE) database. Each sentence is composed of approximately 7–12 words, and each list comprises 10 sentences with an average of 80 words per list. All the words in the sentences were scored for correctness. Two lists for each test condition were used and the scores were averaged. Three conditions were tested for each test: speech in quiet, speech in 10 dB signal to noise ratio (SNR), and speech in 5 dB SNR. Noise type used in all tests was a stationary speech shaped noise created from an average long-term speech spectrum of IEEE sentences.

Electric-only results: Nine subjects were tested for the electric-only condition. Four subjects were tested in the bilateral mode whereas five subjects were tested in the unilateral mode. Figure below presents the percentage correct mean scores of nine subjects as a function of SNR (quiet, SNR = 10 dB, SNR = 5dB) and processor type (clinical, PDA_Offline, PDA_RT). The error bars represent standard error of the mean (SEM). Repeated-measures analysis of variance (ANOVA) was performed to assess the effect of processor and SNR on the intelligibility scores with an α set at 0.05. Subjects were considered a random (blocked) factor while processor types and SNR values were used as main analysis factors. No statistically significant difference in speech intelligibility was found between the three processor types ($F[2,16] = 1.810$, $p = 0.1955$). The interaction be-tween the processor types and SNR was not significant ($F[4,32] = 1.891$, $p = 0.1361$). There was a significant main effect of SNR on speech intelligibility ($F[2,16] = 85.017$, $p<0.0000$). The post hoc Bonferroni test for pairwise comparisons between the three SNRs indicated that speech intelligibility at 10 dB was higher than 5 dB and speech intelligibility was highest in the quiet presentation condition.

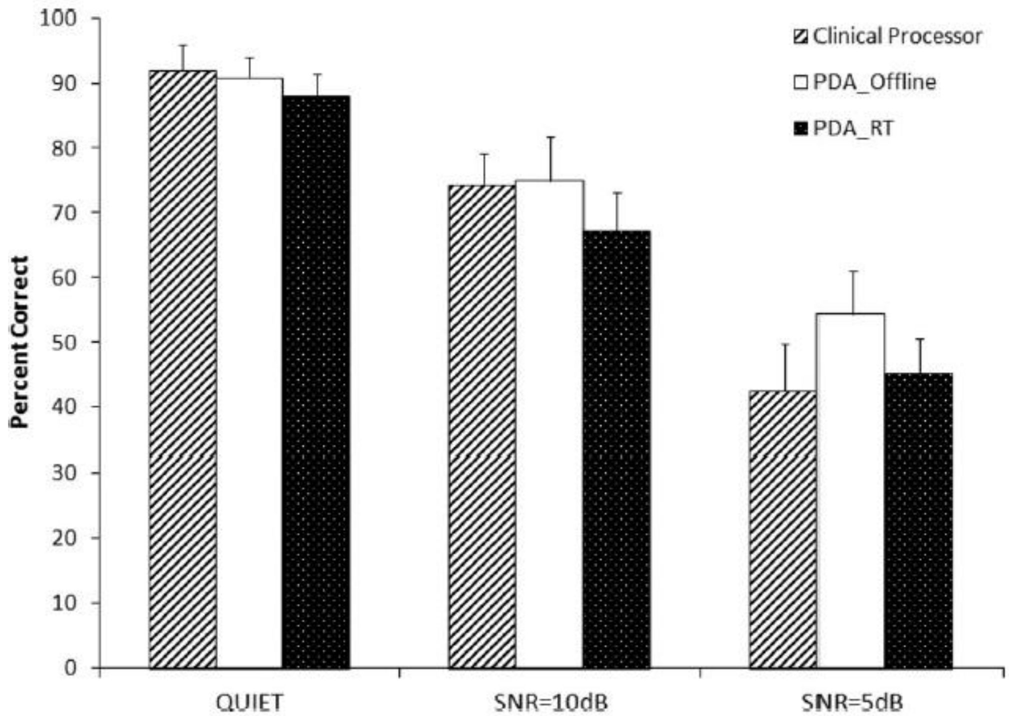


Figure 26: Percentage correct mean speech intelligibility of nine subjects as a function of SNR and processor type. Error bars represent SEM.

Electric Plus Acoustic Stimulation Results: Fig. below displays the mean audiometric thresholds in the non-implanted ear of five bimodal subjects. The mean thresholds at 0.25, 0.5, 0.75, 1.0, 2.0, 3.0, 4.0 and 6.0 kHz were 59, 64, 70, 76, 85, 85, 83, and 87 dB HL, respectively.

Mean speech intelligibility for the EAS condition as functions of SNR and processor type are presented in Fig. below. Repeated measures ANOVA was performed by keeping hearing modalities (acoustic alone, electric alone, and combined EAS), SNRs, and processor types as analysis factors. No statistically significant difference was found between the three processor types ($F[2,6] = 1.968$, $p = 0.2201$). Estimated marginal mean speech intelligibility for acoustic alone (A_alone), electric alone (E_alone), and EAS for the three processor types are shown in Fig. 12. There was a significant effect of hearing modality ($F[2,6] = 6.492$, $p = 0.0316$) on speech intelligibility. The post hoc Bonferroni test for pairwise comparisons between the three hearing modalities indicated that E_alone speech intelligibility was higher than A_alone, and EAS speech intelligibility was significantly higher than E_alone and A_alone speech intelligibilities. Significant interaction between hearing modalities and SNRs was found ($F[4,12] = 5.259$, $p = 0.0111$). The post hoc Bonferroni test for pairwise comparisons ($\alpha = 0.0055$) indicated that speech intelligibility at quiet and 5 dB SNR was significantly different for all three hearing modalities. The results from this extensive evaluation served as a proof for the validity of the platform for CI and HA experiments and later encouraged researchers to use the platform for their future work with E_alone and EAS.

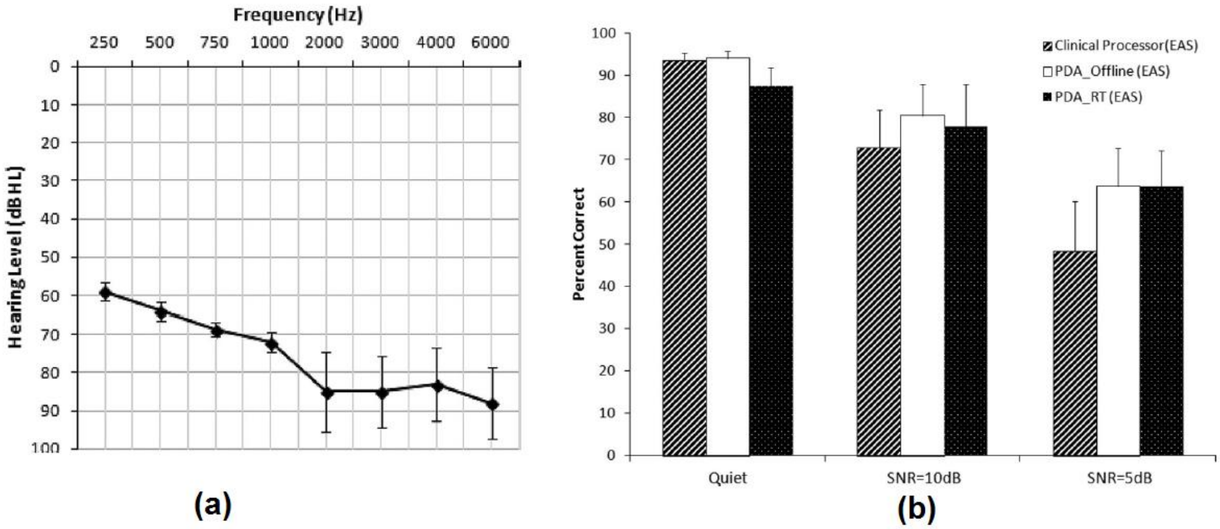


Figure 27: (a) Mean audiogram of five bimodal subjects for the hearing-aid ear. (b) Percentage correct mean speech intelligibility for EAS condition of five bimodal subjects as a function of SNR and processor type. Error bars represent SEM.

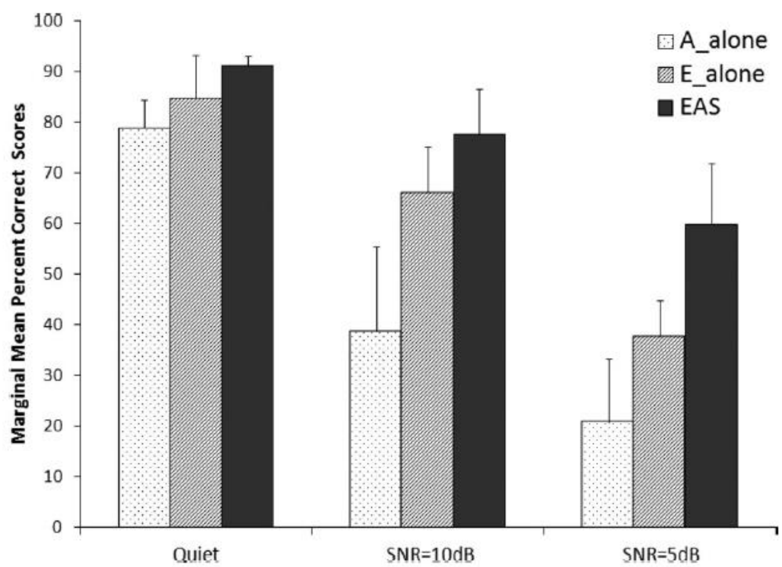


Figure 28: Estimated marginal mean intelligibility scores for acoustic alone (A_alone), electric alone (E_alone), and EAS for the three processor types. Error bars represent SEM.

No adverse effects were observed during testing with the PDA Device

Evaluations with CCi-MOBILE platform – to verify the functionality and safety with human subjects

The latest version of the platform (i.e., with Android interface) also underwent rigorous testing to ensure that it met all safety criteria of a clinical processor. The output levels at the electrodes were verified extensively using 1) an implant emulator and 2) a decoder and implant emulator tool box, from Cochlear Corporation. Electrodograms produced by each processor were compared to ensure similarity for the same input.

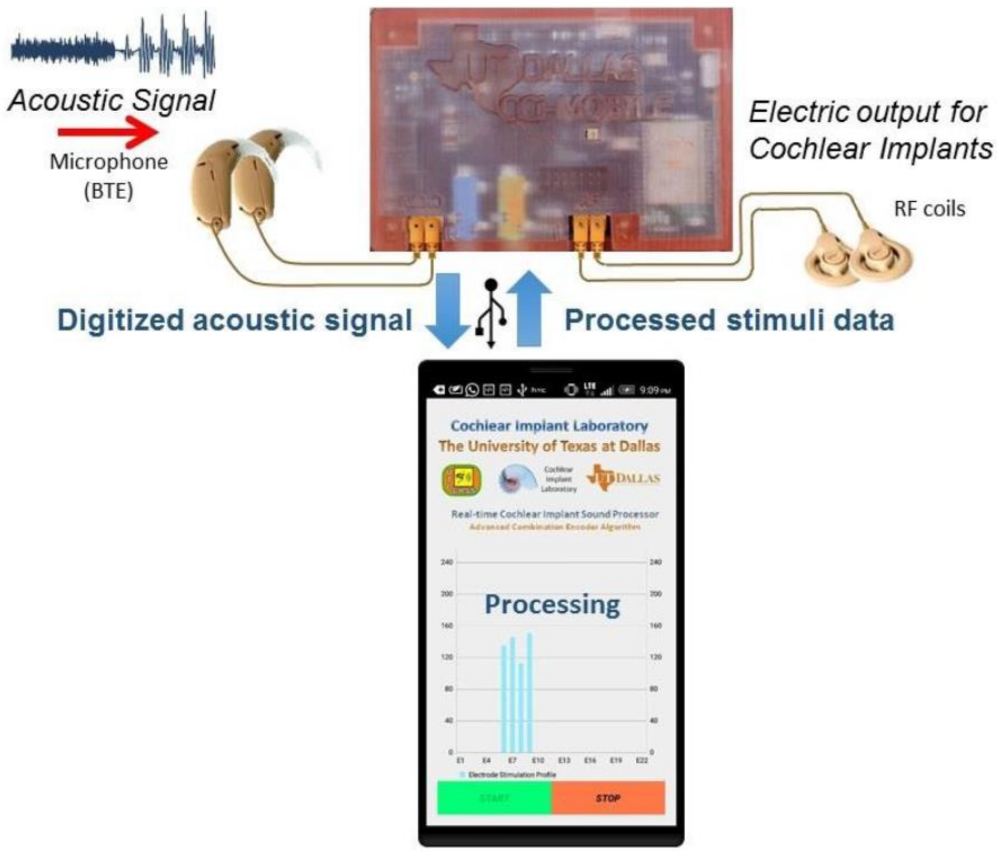


***Figure 29: CCi-MOBILE with Android phone used to assess the efficacy of the research platform.** The acoustic signal is first acquired from BTE unit and is sampled digitally by an onboard stereo codec. The sampled signal is transmitted to the computing platform (Android phone) via USB. The computing platform receives packets of stereo acoustic data, and processes them through a sound coding strategy on a frame-by-frame basis. The processing generates a set of stimulation sequence which is sent back to the interface board where it is encoded (using the transmission protocols of the CI device), and is finally streamed to the implant for stimulation. This process is repeated in real-time.*

In order to assess the efficacy of the research interface, subjective evaluations were carried out with 8 CI users. The aim of the study was to evaluate speech recognition performance of CI users with the CCi-MOBILE research interface and compare performance with their clinical processor. Eight post-lingually deafened adult CI users participated in this study. The assessment of speech recognition was accomplished with AzBio and IEEE sentences presented in quiet, 10dB, and 5dB signal-to-noise ratios as well as with CNC words/phonemes. Study participants were tested in free-field, both with their clinical processor and the research platform. Both devices were programmed with standard ACE sound coding strategy. On all measures of test material, our custom-built mobile research interface produced equivalent performance levels (μ=54.348±6.163) to each individual's clinical processors (μ=52.276±6.318). Repeated Measures Analysis of Variance (ANOVA) revealed no statistically significant difference between the two processor types. The results from this study indicate that performance levels with the research platform are comparable to the clinical processor, and therefore able to accurately duplicate the user's existing clinical configuration. This result suggests great potential for conducting reliable speech assessments in future studies with the CCi-MOBILE research platform.

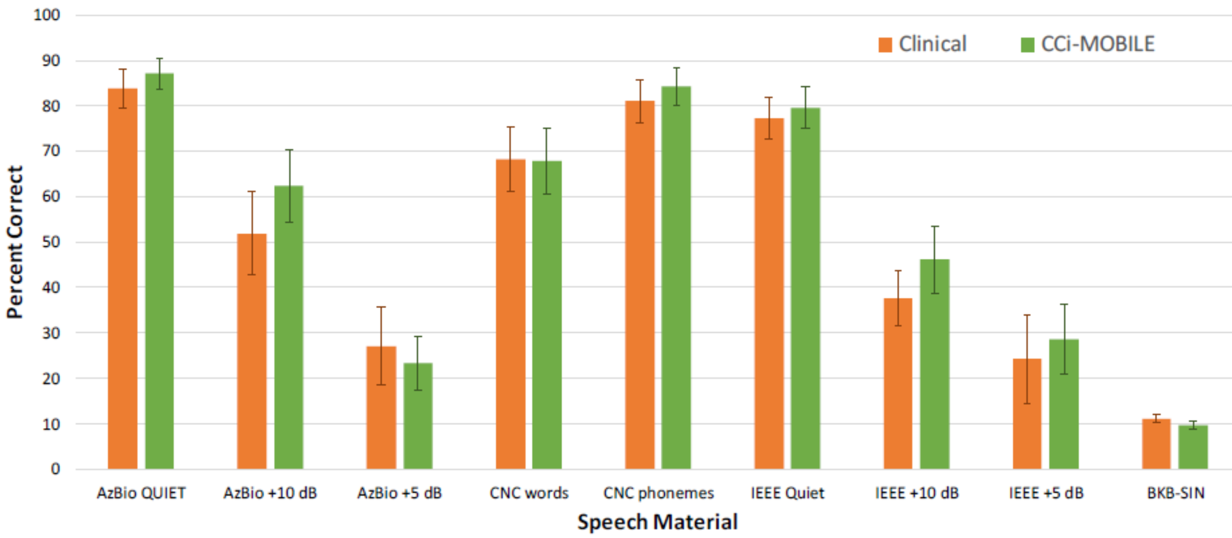


Figure 30: Percentage correct mean speech recognitions scores with clinical processor and CCi-MOBILE research platform. Error bars represent SEM. N = 8.

No adverse effects were observed during testing with the CCi-MOBILE Platform

[Experiments and studies undertaken with the research platform]

The CiPDA and CCiMOBILE research platforms have been used in at least 14 different studies with a total of 97 CI subjects. Following table summarizes the aims of experiments.

	Experiment details	Number of Participants
1.	**Study Title:** Synchronizing Bilateral Cochlear Implants: Preliminary Findings Using the UTD Cochlear Implant PDA (CiPDA) Research Platform in the Free-field. **Objectives:** To evaluate spatial hearing abilities in Bilateral cochlear implant users listening with the ciPDA device. **Reference:** H. Jones, A. Kan and R. Litovsky, "Synchronizing Bilateral Cochlear Implants: Preliminary Findings Using the UTD Cochlear Implant PDA (CiPDA) Research Platform in the Free-field," in 38th Annual mid-winter meeting of Association for Research in Otolaryngology, Baltimore, 2015.	**10**
2.	**Study Title:** Evaluation and analysis of whispered speech for cochlear implant users: Gender identification and intelligibility **Objectives:** To investigate the degree to which whispered speech impacts speech impacts perception and gender identification in cochlear implant (CI) users. **Reference:** O. Hazrati, H. Ali and J. H.L. Hansen, "Evaluation and analysis of whispered speech for cochlear implant users: gender identification and intelligibility," Journal of Acoustical Society of America, vol. 138, pp. 74-79, 2015.	**6**
3.	**Study Title:** Simultaneous suppression of noise and reverberation in cochlear implants using a ratio masking strategy **Objectives:** Assess speech perception (speech intelligibitly) with a new ratio masking strategy to suppress noise and reverberation. **Reference:** O. Hazrati, S. O. Sadjadi, P. C. Loizou and J. H. L.Hansen, "Simultaneous Suppression of Noise and Reverberation in Cochlear Implants Using a Ratio Masking Strategy," Journal of Acoustical Society of America, vol. 134, no. 5, pp. 3759-3765, 2013.	**7**
4.	**Study Title:** Blind binary masking for reverberation suppression in cochlear implants. **Objectives:** Assess benefits with a novel monaural binary time-frequency (T-F) masking technique for suppressing reverberation for cochlear implant users. **Reference:** O. Hazrati, J. Lee and P. C. Loizou, "Blind Binary Masking for Reverberation Suppression in Cochlear Implants," Journal of Acoustical Society of America, vol. 133, no. 3, pp. 1607-1614, 2013.	**6**

5.	**Study Title:** Noise Modeling in Reverberation: A Comparative Study of Speech Intelligibility in Cochlear Implant Users **Objectives:** To investigate the speech intelligibility of CI users in reverberant noisy (RN) and noisy reverberant (NR) listening environments **Reference:** H. Ali, O. Hazrati, N. Srinivasan and P. C. Loizou, "Noise Modeling in Reverberation: A Comparative Study of Speech Intelligibility in Cochlear Implant Users," in Sigma Xi Southwest Regional Research Conference, Dallas, TX, 2013.	**6**
6.	**Study Title:** Effect of consonant-vowel boundary to speech perception in cochlear implants **Objectives:** To study the relative contributions of masked consonant and vowel sounds on speech perception with cochlear implant devices **Reference:** N. Srinivasan, H. Ali, E. A. Tobey and P. C. Loizou, "Effect of consonant-vowel boundary to speech perception in cochlear implants," in Conference on Implantable Auditory Prostheses, Lake Tahoe, 2013.	**8**
7.	**Study Title:** Temporal modulation transfer functions for low modulation frequencies in adult cochlear implant listeners **Objectives:** to derive relationships between modulation detection thresholds as a function of low-level modulation frequencies in cochlear implant recipients **Reference:** N. Srinivasan, H. Ali, E. A. Tobey and P. C. Loizou, "Temporal modulation transfer functions for low modulation frequencies in adult cochlear implant listeners," submitted: International Journal of Audiology.	**10**
8.	**Study Title:** Image-Guided Frequency-Place Mapping in Cochlear Implants **Objectives:** Assess the benefits of a novel image guided programming strategy for cochlear implant users. **Reference:** H. Ali, J. H. Noble, R. H. Gifford, R. F. Labadie, B. M. Dawant, J. H.L. Hansen and E. A. Tobey, "Image-Guided Frequency-Place Mapping in Cochlear Implants," in Conference on Implantable and Auditory Prostheses, Lake Tahoe, 2015.	**2**
9.	**Study Title:** Improving channel selection of sound coding algorithms in cochlear implants **Objectives:** To test the effectiveness of a novel channel selection algorithm for cochlear implant sound processors **Reference:** H. Ali, F. Hong, J. H. L. Hansen and E. A. Tobey, "Improving channel selection of sound coding algorithms in cochlear implants," in International Conference on Acoustics, Speech, and Signal Processing, Florence, 2014.	**3**

10.	**Study Title:** Evaluation of a PDA-based Research Platform for Cochlear Implants **Objectives:** To test the efficacy of Ci-PDA device for studies with cochlear implant users **Reference:** H. Ali, A. Lobo and P. C. Loizou, "Design and Evaluation of a PDA-based Research Platform for Cochlear Implants," IEEE Transactions on Biomedical Engineering, vol. 60, no. 11, pp. 3060-3073, 2013.	**10**
11.	**Study Title:** Bimodal Cochlear Implants: The Role of Acoustic Signal Level in Determining Speech Perception Benefit **Objectives:** To determine, for bimodal cochlear implant (CI) patients, i.e., patients with low-frequency hearing in the ear contralateral to the implant, how speech understanding varies as a function of the difference in level between the CI signal and the acoustic signal. **Reference:** M. F. Dorman, P. C. Loizou, S. Wang, T. Zhang, A. Spahr, L. Loiselle and S. Cook, "Bimodal Cochlear Implants: The Role of Acoustic Signal Level in Determining Speech Perception Benefit," Audiology & Neurotology, vol. 19, no. 4, pp. 234-238, 2014.	**5**
12.	**Study Title:** Binaural Enhancement for Bilateral Cochlear Implant Users **Objectives:** To investigate the feasibility of a method for enhancing the binaural cues available to BCI users **Reference:** C. A. Brown, "Binaural Enhancement for Bilateral Cochlear Implant Users," Ear and Hearing, vol. 35, no. 5, pp. 580-584, 2014.	**8**
13.	**Study Title:** "Subjective evaluation with UT-Dallas research interface for cochlear implant users **Objectives:** To test the efficacy of CCi-MOBILE platform for studies with cochlear implant users **Reference:** H. Ali, S. Ammula, J.H.L. Hansen, "Subjective evaluation with UT-Dallas research interface for cochlear implant users," Conf. on Implantable Auditory Prostheses, CIAP-2017, pp. 211, Lake Tahoe, CA, July 16-21, 2017	**8**
14.	**Study Title:** Improving channel selection criteria in n-of-m strategies for cochlear implant sound coding strategies **Objectives:** To test the effectiveness of a novel channel selection algorithm for cochlear implant sound processors. **Reference**: J.N. Saba, H. Ali, J.H.L. Hansen, "Improving channel selection criteria in n-of-m strategies for cochlear implant sound coding strategies," Conf. on	**8**

Implantable Auditory Prostheses, CIAP-2017, pp. 227, Lake Tahoe, CA, July 16-21, 2017

[Speech enhancement algorithms for cochlear implants]

Noise, reverberation, and combination of both interferences continues to be a huge challenge for CI users as speech recognition performance by CI users significantly reduces even at minor interference levels (at which normal hearing individuals can understand speech at ceiling levels). Noise reduction and speech enhancement methods can reduce the negative effects of interference and provide better signal representation for improved speech recognition. A major effort was devoted to addressing speech recognition challenges with cochlear implant devices in everyday noisy environments. In this regard, several novel signal processing strategies were developed specifically for cochlear implant users that aimed at suppressing noise and reverberation, enhancing speech signal for better intelligibility of speech, and improving the overall listening quality. The sound processing schemes developed throughout the lifecycle of this project were formulated in a way that they can be easily integrated in commercial sound processors. The improvements in speech recognition performance observed with the proposed approaches (discussed below) can potentially help hundreds of thousands of CI users in better speech understanding and quality in adverse listening conditions.

1. Reverberation and noise suppression algorithms for cochlear implant users

Although reverberation does not significantly affect speech quality and intelligibility for normal hearing (NH) listeners, it severely degrades word identification performance in hearing impaired (e.g., cochlear implant (CI) users) as well as elderly people.

During this project, word recognition performance of cochlear implantees in reverberant conditions was extensively studied and new reverberation suppression strategies were proposed in order to improve intelligibility. Effectiveness of the proposed techniques was first evaluated in ideal scenarios, where access to an anechoic signal is assumed, to determine an upper-bound on performance. Formal listening experiments indicated large speech intelligibility gains in low to highly reverberant environments which motivated development of non-ideal suppression methods. Two novel channel selection criteria to blindly (non-ideally) suppress the reverberation in CI devices were developed. These techniques lead to significant improvements for CI subjects in word identification performance in reverberation and can easily be integrated into current CI devices.

In addition to reverberation, the presence of background noise can also degrade word identification scores of CI users. Noise masks speech in a different manner compared to reverberation, and when added to reverberation, further elevates the confusion for listeners compared to reverberation/noise alone conditions. The combined effects of noise and reverberation are quite detrimental to speech intelligibility especially for CI users with only a limited number of stimulating electrodes in their devices. Several studies were conducted to assess the performance of CI listeners in more realistic environments where reverberation and noise co-exist. Following subsections provide further details on specific algorithms and human studies conducted with unilateral CI users as well as outcomes.

A channel-selection criterion for suppressing reverberation in cochlear implants: The first step in this process was to quantify the extent to which reverberation affects speech intelligibility by CI listeners. In this regard an experiment was conducted to assess CI users' performance using Institute of Electrical and Electronics Engineers (IEEE) sentences corrupted with varying degrees of reverberation. Reverberation times of 0.30, 0.60, 0.80, and 1.0 s were used. Results indicated that for all subjects tested, speech intelligibility decreased exponentially with an increase in reverberation time. A decaying-exponential model provided an excellent fit to the data. In order to combat the negative effects of reverberation a speech coding strategy was devised for reverberation suppression using a channel-selection criterion based on the signal-to-reverberant ratio (SRR) of individual frequency channels. The SRR reflects implicitly the ratio of the energies of the signal originating from the early (and direct) reflections and the signal originating from the late reflections. Channels with SRR larger than a preset threshold were selected, while channels with SRR smaller than the threshold were zeroed out. Results in a highly reverberant scenario indicated that the proposed strategy led to substantial gains (over 60 percentage points) in speech intelligibility over the subjects' daily strategy. Further analysis indicated that the proposed channel-selection criterion reduced the temporal envelope smearing effects introduced by reverberation and also diminishes the self-masking effects responsible for flattened formants.

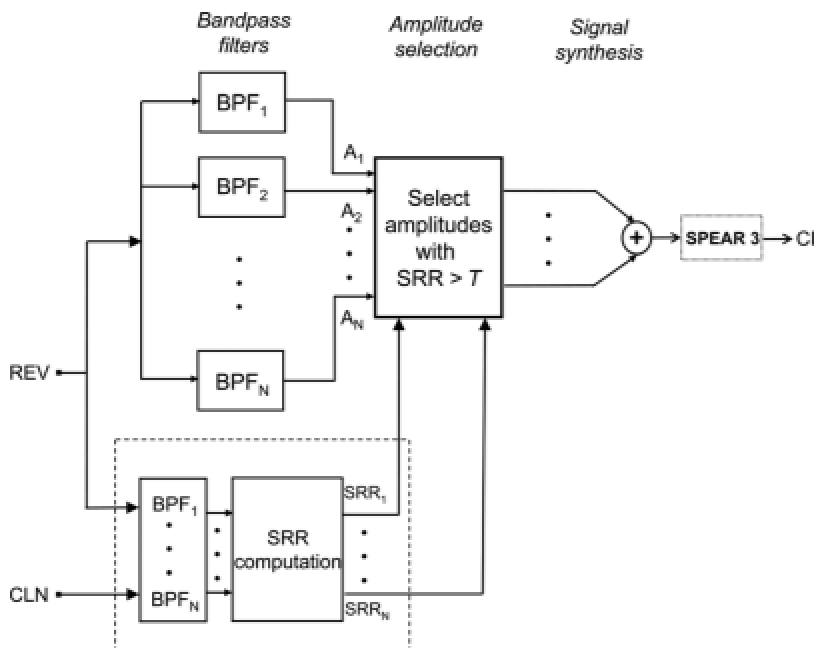
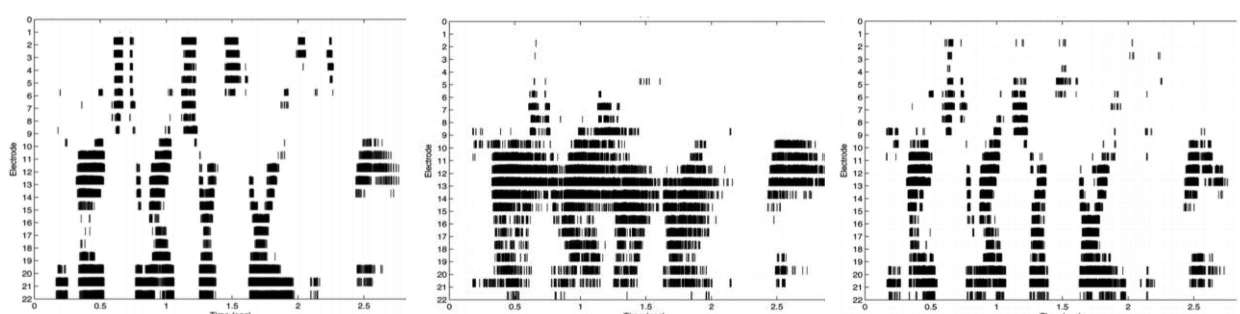


Figure 31: Block diagram of the IRM strategy.



(a) (b) (c)

***Figure 32: Stimulus output patterns (electrodograms) of the IEEE sentence:* "The urge to write short stories is rare" uttered by a male speaker. (a) Electrodogram of unmodified (uncorrupted) sentence processed by the ACE strategy, (b) Electrodogram of the same sentence when corrupted by reverberation equal to RT_{60} = 1.0 s and processed by the ACE strategy, and (c) electrodogram of the reverberant sentence when processed by the IRM speech coding strategy with the threshold set to -5 dB. In each electrodogram, time is shown along the abscissa and the electrode number is shown along the ordinate**

Further details on the algorithm, methods, and evaluations can be found in: Kostas Kokkinakis, Oldooz Hazrati, and Philipos C. Loizou, "A Channel-Selection Criterion for Suppressing Reverberation in Cochlear Implants", J. Acoust. Soc. Am., 129(5), 3221-3232, 2011.

Tackling the Combined Effects of Reverberation and Masking Noise Using Ideal Channel Selection: The above work was expanded by developing and evaluating a new signal processing algorithm for the suppression of the combined effects of reverberation and noise. The proposed algorithm decomposed, on a short-term basis (every 20 ms), the reverberant stimuli into a number of channels and retained only a subset of the channels satisfying a signal-to-reverberant ratio (SRR) criterion. The construction of this criterion assumes access to a priori knowledge of the target (anechoic) signal and the aim of the study was to assess the full potential of the proposed channel-selection algorithm assuming that this criterion could be estimated accurately. Listening tests were conducted with normal-hearing listeners to assess the performance of the proposed algorithm in highly reverberant conditions (T60 = 1.0 s) which included additive noise at 0 and 5 dB SNR.

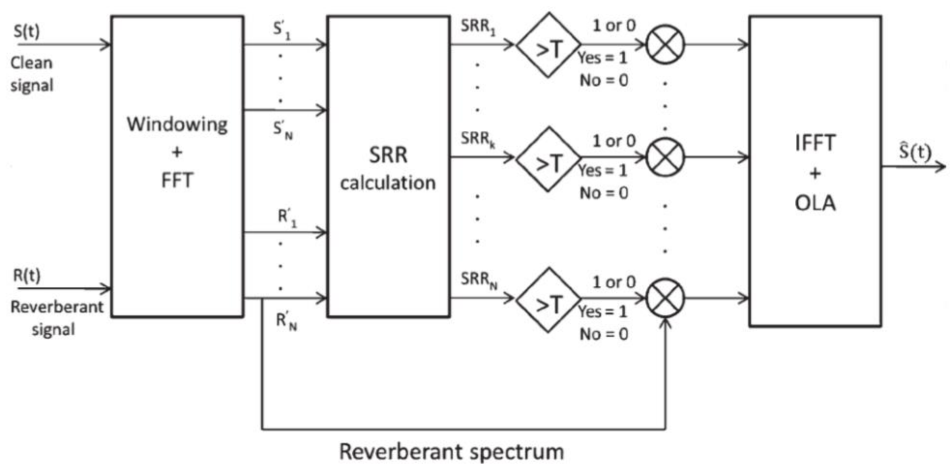


Figure 33: Block diagram of the proposed ideal channel-selection (ICS) algorithm.

A substantial gain in intelligibility was obtained in both reverberant and combined reverberant and noise conditions. The mean intelligibility scores improved by 44 and 33 percentage points at 0 and 5 dB SNR reverberant+noise conditions. Feature analysis of the consonant confusion matrices revealed that the transmission of voicing information was most negatively affected, followed by manner and place of articulation.

The proposed algorithm was found to produce substantial gains in intelligibility, and this benefit was attributed to the ability of the proposed SRR criterion to accurately detect voiced/unvoiced boundaries. Detection of those boundaries is postulated to be critical for better perception of voicing information and manner of articulation.

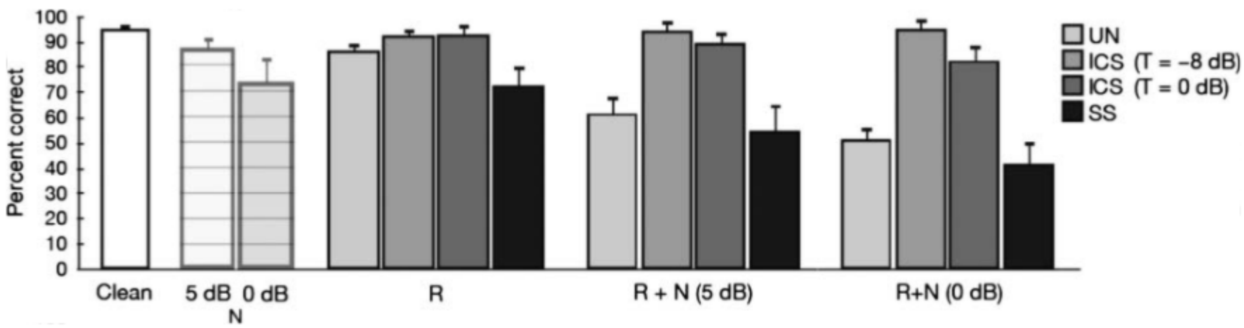


Figure 34: Mean intelligibility scores obtained in the various conditions involving reverberation (R), noise (N), and combined reverberation and noise (R + N). The leftmost bar shows performance obtained in (clean) anechoic conditions. Scores obtained with unprocessed stimuli are labeled as UN, and scores obtained with stimuli processed via the spectral-subtractive algorithm are labeled as SS.

Further details on the algorithm, methods, and evaluations can be found in: Oldooz Hazrati, and Philipos C. Loizou, "Tackling the Combined Effects of Reverberation and Masking Noise using Ideal Channel Selection", J. Speech. Lang. Hear. Res., 55, 500-510, 2012.

Predicting the intelligibility of reverberant speech for cochlear implant listeners with a non-intrusive intelligibility measure: Reverberation is known to reduce the temporal envelope modulations present in the signal and affect the shape of the modulation spectrum. A non-intrusive intelligibility measure for reverberant speech was formulated, motivated by the fact that the area of the modulation spectrum decreases with increasing reverberation. The proposed measure was based on the average modulation area computed across four acoustic frequency bands spanning the signal bandwidth. High correlations (r = 0.98) were observed with sentence intelligibility scores obtained by cochlear implant listeners. Proposed measure outperformed other measures including an intrusive speech-transmission index based measure.

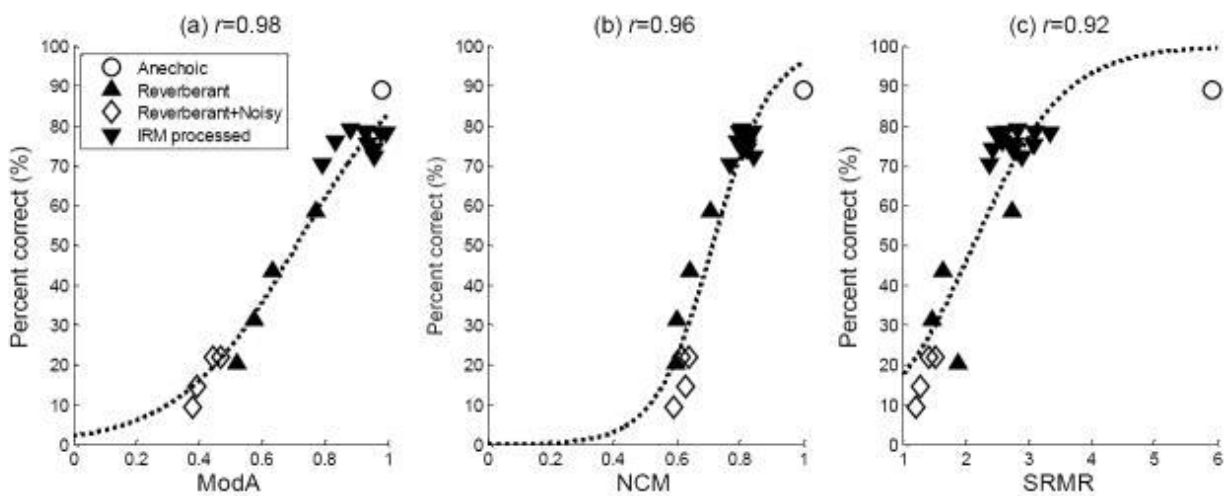


Figure 35: Scatter plots of sentence recognition scores against the: (a) ModA, (b) NCM, and (c) SRMR values. 'IRM' denotes the ideal reverberant mask algorithm.

Further details on the algorithm, methods, and evaluations can be found in: Fei Chen, Oldooz Hazrati, and Philipos C. Loizou, "Predicting the Intelligibility of Reverberant Speech for Cochlear Implant Listeners with a Non-intrusive Intelligibility Measure", Biomedical Signal Processing & Control., 8(3), 311-314, 2013

Simultaneous suppression of noise and reverberation in cochlear implants using a ratio masking strategy: Our experiments indicated that the speech identification task for CI users becomes even more challenging in conditions where both reverberation and noise co-exist as they mask the spectro-temporal cues of speech in a rather complementary fashion. Ideal channel selection (ICS) was found to result in significantly more intelligible speech when applied to the noisy, reverberant, as well as noisy reverberant speech. Further research was conducted to develop more advanced algorithms for reverberation and noise suppression. One such algorithm was a blind single-channel ratio masking strategy that simultaneously suppressed the negative effects of reverberation and noise on speech identification performance for CI users. In this strategy, noise power spectrum was estimated from the non-speech segments of the utterance while reverberation spectral variance was computed as a delayed and scaled version of the reverberant speech spectrum. Based on the estimated noise and reverberation power spectra, a weight between 0 and 1 was assigned to each time-frequency unit to form the final mask. Listening experiments conducted with CI users in two reverberant conditions (T60 = 0.6 and 0.8 s) at a signal-to-noise ratio of 15 dB indicated substantial improvements in speech intelligibility in both reverberant-alone and noisy reverberant conditions.

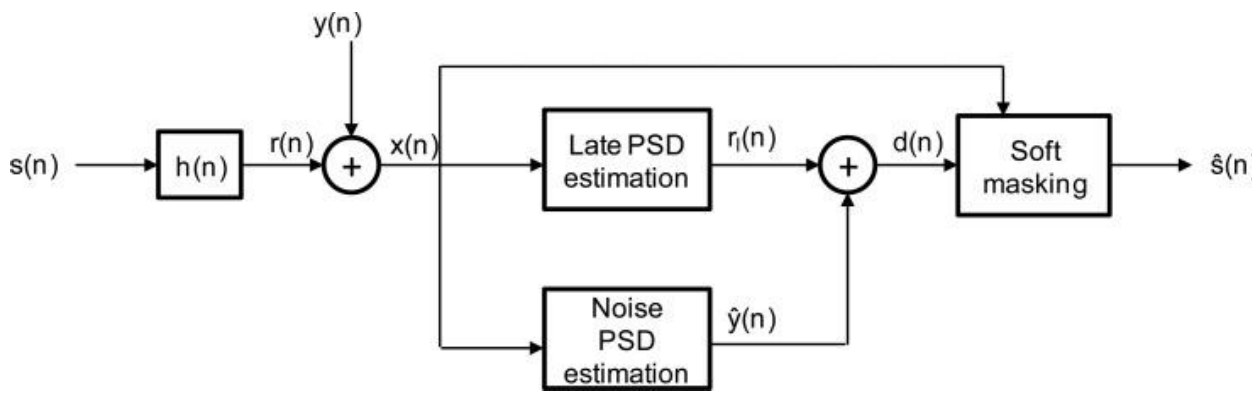


Figure 36: Block diagram of the proposed noisy reverberant speech enhancement algorithm. *Here, $\hat{y}(n)$, $r_l(n)$, and d(n) denote the estimated noise PSD, late reverberation PSD, and the total estimated distortions, respectively.*

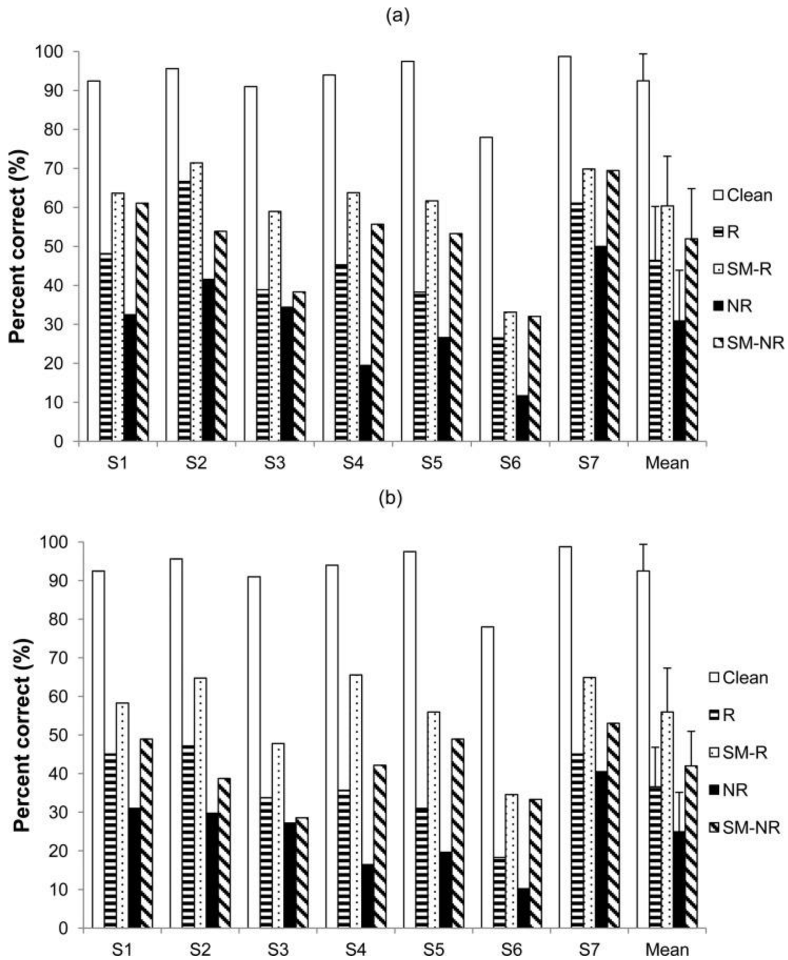


Figure 37: Individual percent correct scores of seven CI users tested on IEEE sentences using unprocessed and soft-mask processed reverberant and noisy reverberant acoustic inputs (RSNR = 15 dB), (a) T60 = 0.6 s and (b) T60 = 0.8 s. "Clean," "R," "NR," and "SM" stand for anechoic clean, reverberant, noisy reverberant, and soft-mask processed conditions, respectively. Error bars indicate standard deviations.

Further details on the algorithm, methods, and evaluations can be found in: Oldooz Hazrati, Seyed Omid Sadjadi, Philipos C. Loizou, and John H.L. Hansen, "Simultaneous Suppression of Noise and Reverberation in Cochlear Implants Using a Ratio Masking Strategy", J. Acoust. Soc. Am., 134, 2013

Reverberation suppression in cochlear implants using a blind channel-selection strategy: The ideal reverberant mask (IRM), a binary mask for reverberation suppression which is computed using signal-to-reverberant ratio, was found to yield substantial intelligibility gains for CI users even in highly reverberant environments (e.g., T60 = 1.0 s). Motivated by the intelligibility improvements obtained from IRM, a monaural blind channel-selection criterion for reverberation suppression was proposed. The proposed channel-selection strategy is blind, meaning that prior knowledge of neither the room impulse response nor the anechoic signal is required. By the use of a residual signal obtained from linear prediction analysis of the reverberant signal, the residual-to-reverberant ratio (RRR) of individual frequency channels was employed as the channel-selection criterion. In each frame, the channels with RRR less than an adaptive threshold were retained while the rest were zeroed out. Performance of the proposed strategy was evaluated via intelligibility listening tests conducted with CI users in simulated rooms with two reverberation times of 0.6 and 0.8 s. The results indicated significant intelligibility improvements in both reverberant conditions (over 30 and 40 percentage points in

T60 = 0.6 and 0.8 s, respectively). The improvement was observed to be comparable to that obtained with the IRM strategy.

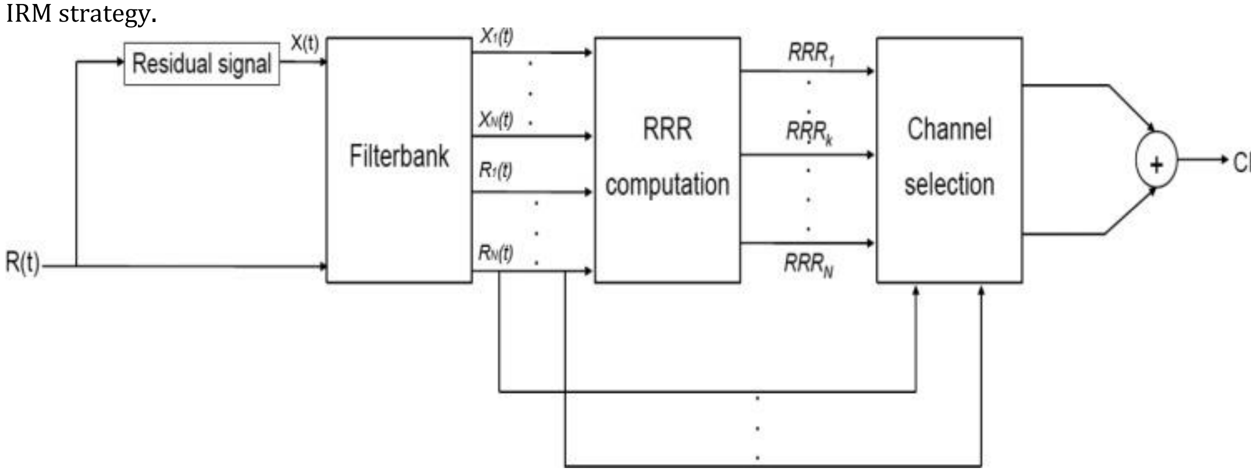


Figure 38: Block diagram of the proposed channel-selection algorithm

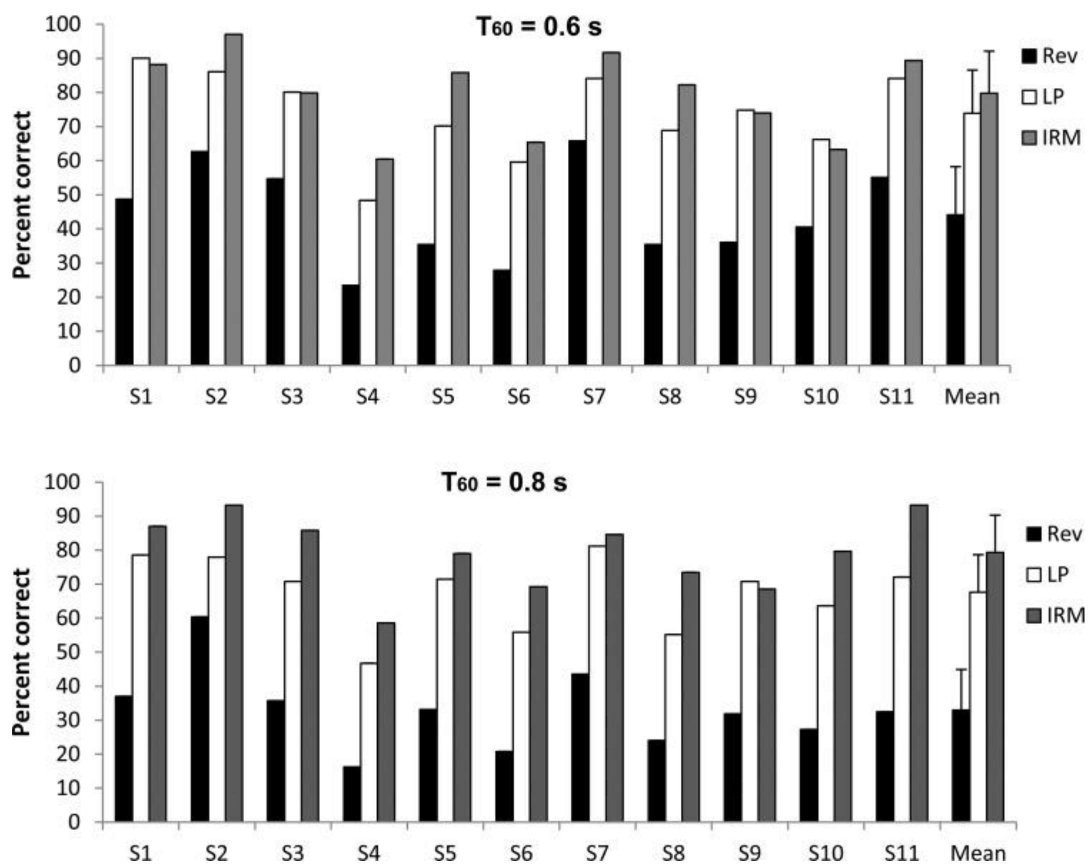


Figure 39: Individual percent correct scores of eleven CI users tested on IEEE sentences using anechoic (unprocessed) acoustic inputs recorded in reverberation (black bars) and reverberant acoustic inputs processed with the LP channel-selection criterion (white bars) and IRM strategy (gray bars), (a) T_{60} = 0.6 s and (b) T_{60} = 0.8 s. Error bars indicate standard deviations.

Further details on the algorithm, methods, and evaluations can be found in: Oldooz Hazrati, and Philipos C. Loizou, "Reverberation Suppression in Cochlear Implants Using a Blind Channel-Selection Strategy", J. Acoust. Soc. Am., 133(6), 4188-4196, 2013.

Blind binary masking for reverberation suppression in cochlear implants: Another effort in reverberation mitigation was the development of a monaural binary time-frequency (T-F) masking technique. In this technique, the mask was estimated for each T-F unit by extracting a variance-based feature from the reverberant signal and comparing it against an adaptive threshold. Performance of the estimated binary mask was evaluated in three moderate to relatively high reverberant conditions (T60 = 0.3, 0.6, and 0.8 s) using intelligibility listening tests with cochlear implant users. Results indicated that the proposed T-F masking technique yielded significant improvements in intelligibility of reverberant speech even in relatively high reverberant conditions (T60 = 0.8 s). The improvement was hypothesized to result from the recovery of the vowel/consonant boundaries, which are severely smeared in reverberation.

Further details on the algorithm, methods, and evaluations can be found in: Oldooz Hazrati, Jaewook Lee, and Philipos C. Loizou, "Blind Binary Masking for Reverberation Suppression in Cochlear Implants", J. Acoust. Soc. Am., 133(3), 1607-1614, 2013.

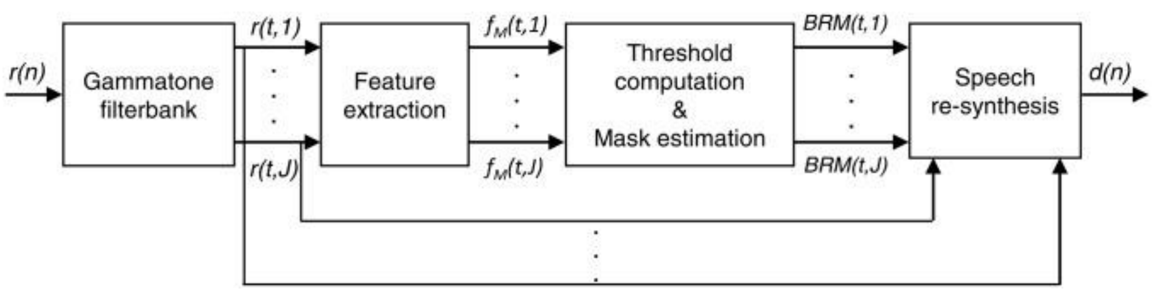


Figure 40: Block diagram of the proposed binary mask estimation technique for de-reverberation

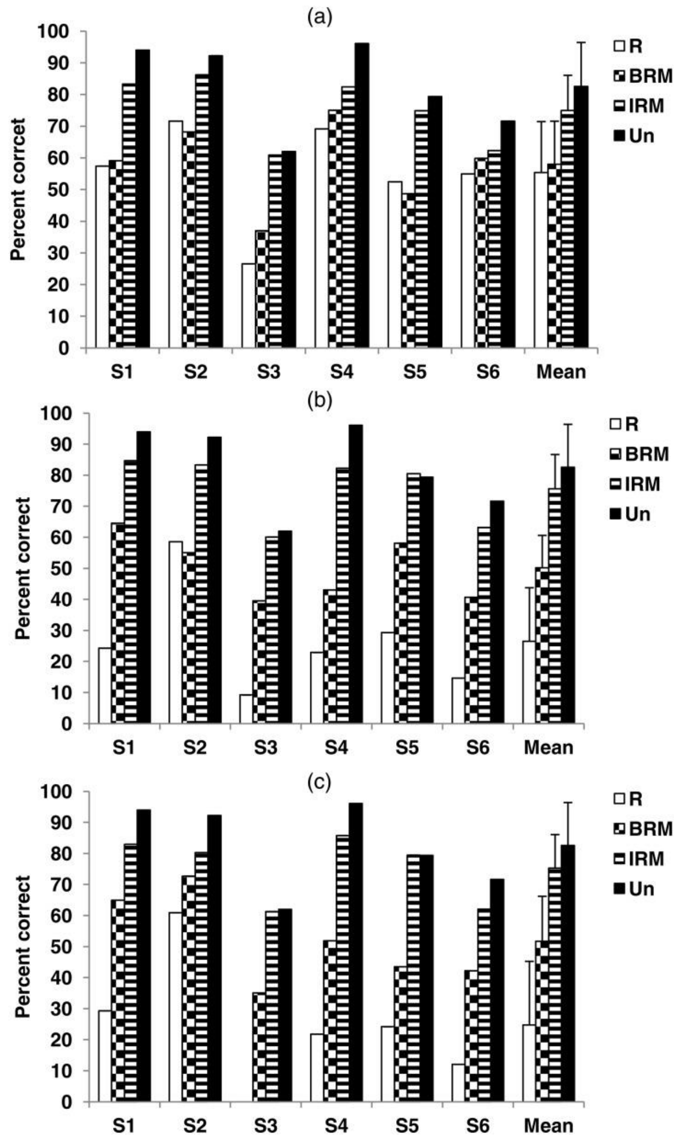


***Figure 41: Individual as well as average intelligibility scores of six CI users** in (a) T_{60} = 0.3 s, (b) T_{60} = 0.6 s, and (c) T_{60} = 0.8 s. "R," "BRM," "IRM," and "Un" stand for reverberant signals, BRM-processed, IRM-processed, and unprocesed signals, respectively. Error bars indicate standard deviations.*

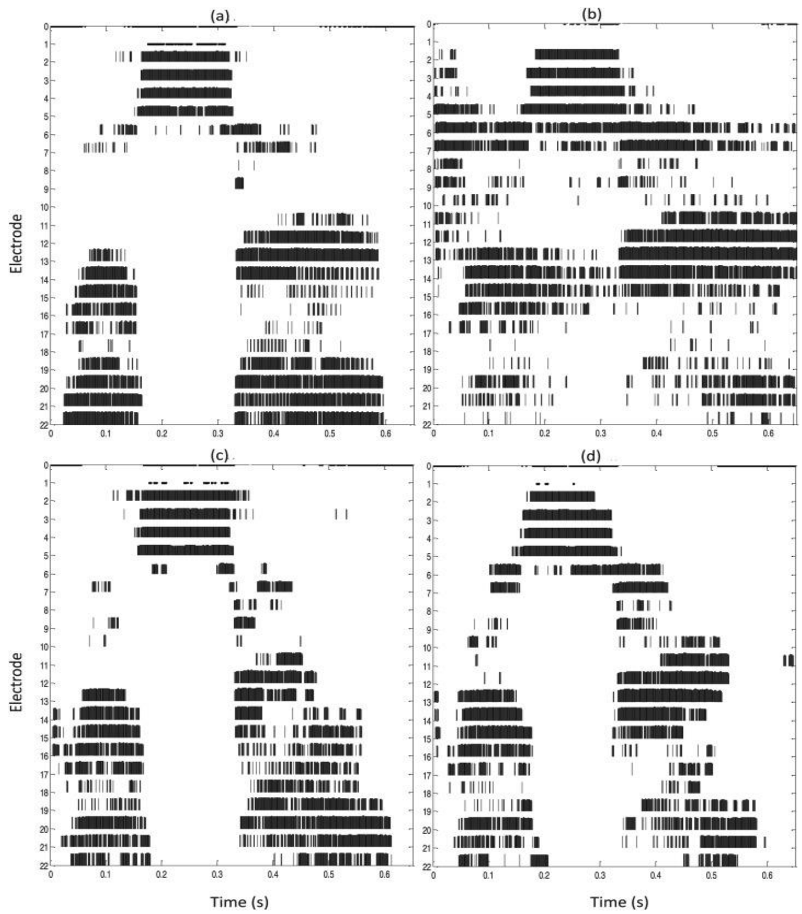


Figure 42: Stimulus output patterns (electrodograms) of the words: "will serve" from the IEEE sentence "The stitch will serve but needs to be shortened," for (a) clean, (b) reverberant (T_{60} = 0.6 s), (c) IRM-processed, and (d) BRM-processed reverberant signals.

2. Speech enhancement strategies for improved speech intelligibility and quality for cochlear implant listeners

One of the reasons speech perception with cochlear implant devices is suboptimal in noise is because of the low spectral resolution provided by CI encoding strategies is insufficient to distinguish speech components from noise. In order to mitigate the effects of noise, several new speech enhancement solutions were proposed and evaluated that aimed at improving speech intelligibility for CI recipients in diverse noise types. The premise of the algorithms was to use harmonic structure of speech as the frequency domain cues to estimate the degraded noise. This was inspired by the fact that speech energy is primarily carried in the harmonic structure located at multiple integer harmonics of the fundamental frequency. The proposed speech enhancement methods are based on harmonic structure estimation combined with a traditional statistical based leveraged solution. Specifically, within the scope of this project and grant, investigations were carried out on robust fundamental frequency estimation in noise, along with integrating such novel techniques to improve harmonic based speech enhancement in both stationary and non-stationary noise scenarios.

Noise-robust F0 estimation from temporal harmonic structures in local TF segments: Fundamental frequency (F0) is one of the most important characteristics of human speech which represents the vibration rate of the vocal cords during speech production. A robust F0 estimation system has potential to facilitate many speech signal processing areas, such as speech source separation, emotion recognition, speaker/language identification, etc. The accuracy of pitch estimation decreases dramatically in the low SNR conditions, due to the fact that both periodic cues in time domain and harmonic structures in frequency domain are distorted. In order to reduce the noise affect, we took advantage of the sparsity of speech signal that high energy of speech is only distributed in a certain number of TF segments. These TF segments will have high SNR levels compared to the overall average SNR of the full-band signal where speech is the dominated element. By performing pitch estimation in these TF segments, it is possible to improve the overall accuracy. In this regard, a robust F0 estimation algorithm was developed. The proposed algorithm included two stages. First, a series of F0 candidate contours were estimated from different TF segments. Second, F0 tracking was performed across the overall TF plane to select the best F0. The harmonic structures in these high SNR TF segments provide a more reliable source for F0 estimation in noise. An algorithm overview shown the F0 estimation is presented in Fig. 3.2. Experiments and results have shown that the proposed algorithm substantially outperforms the compared state-of-the-art methods in terms of pitch estimation accuracy.

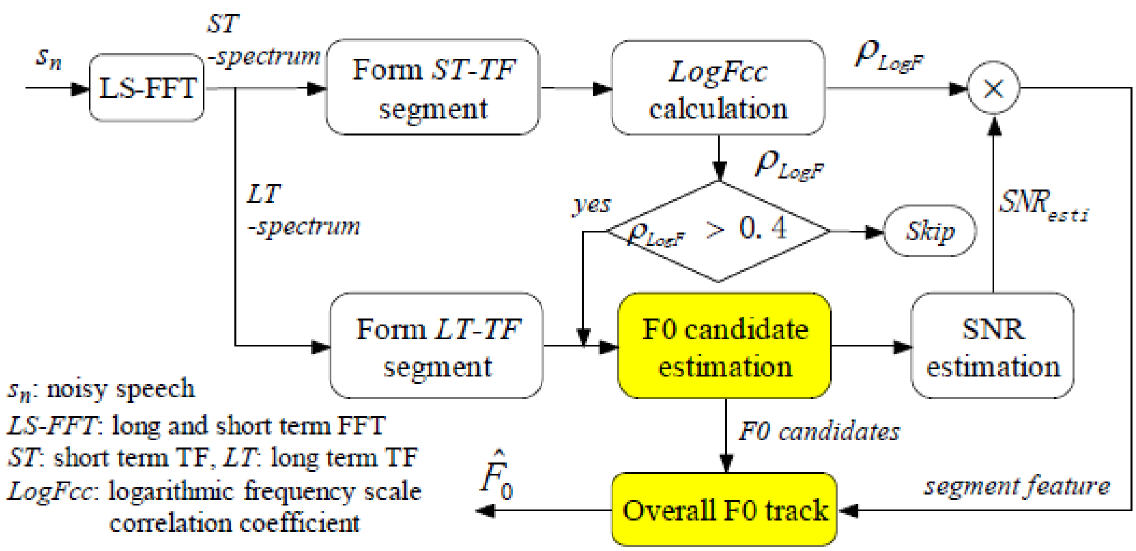


Figure 43: Signal flow in the developed F0 estimation algorithm

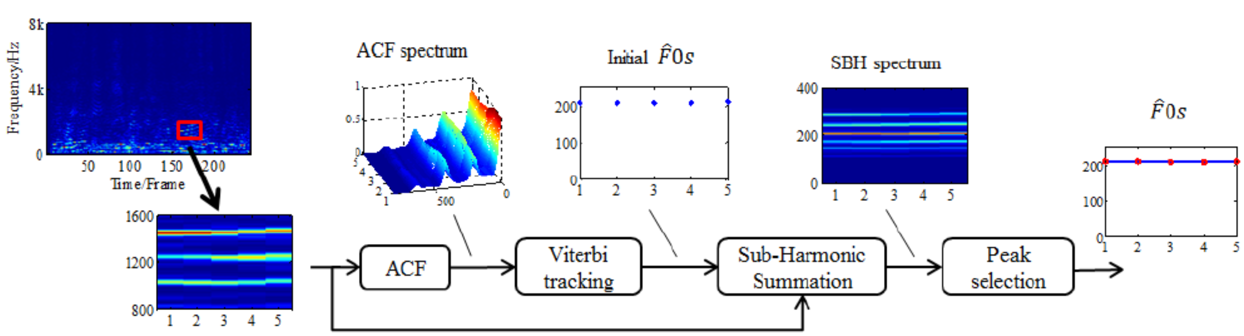


Figure 44: Overview of F0 candidate estimation for the developed F0 estimation algorithm.

Robust harmonic features for neural network classification-based pitch estimation: As pointed out above pitch estimation in diverse naturalistic audio streams remains a challenge for speech processing and spoken language technology. In order to develop robust speech enhancement algorithms that are based on pitch and harmonic structure of speech signal, we investigate the use of robust harmonic features for classification-based pitch estimation. A robust pitch estimation algorithm was developed which was composed of two stages: pitch candidate generation and target pitch selection. Based on energy intensity and spectral envelope shape, five types of robust harmonic features were proposed to reflect pitch associated harmonic structure. A neural network was adopted for modeling the relationship between input harmonic features and output pitch salience for each specific pitch candidate. In the test stage, each pitch candidate was assessed with an output salience which indicated the potential as a true pitch value, based on its input feature vector processed through the neural network. Finally, according to the temporal continuity of pitch values, pitch contour tracking was performed using a hidden Markov model (HMM) and the Viterbi algorithm was used for HMM decoding. Experimental results show that the proposed algorithm outperformed several state-of-the-art pitch estimation methods in terms of accuracy in both high and low levels of additive noise.

Pitch estimation is formulated as a pitch candidates classification problem in this study that advanced machine learning approach is able to be used to improve the accuracy in noise. Pitch information is implicitly contained in the harmonic structures of voiced speech sections. Features extracted from harmonic structure can describe the characteristics of the corresponding pitch candidate. These harmonic features mapped the pitch candidates into a more separable space for the classification step. The analysis of the harmonic features is presented in Fig. 45 and Fig. 46. The proposed method consisted of two processing steps. In the first step, pitch candidates were generated from the original noisy speech spectrum as well as SBH spectrum. In the second step, pitch candidate classification was performed based on a neural network solution using multi-dimensional pitch related robust harmonic features. Five robust features are proposed to provide complementary information for neural network based pitch classification. Finally, by applying pitch temporal continuity constraints, the resulting pitch tracking was based on an HMM to select the optimal and smoothed pitch contours. Experimental results demonstrated that the proposed HarFeature algorithm yielded substantially better performance (lower GPE) than the compared state-of-the-art algorithms across various types and levels of noise, which is shown in Fig 47 and Fig. 48.

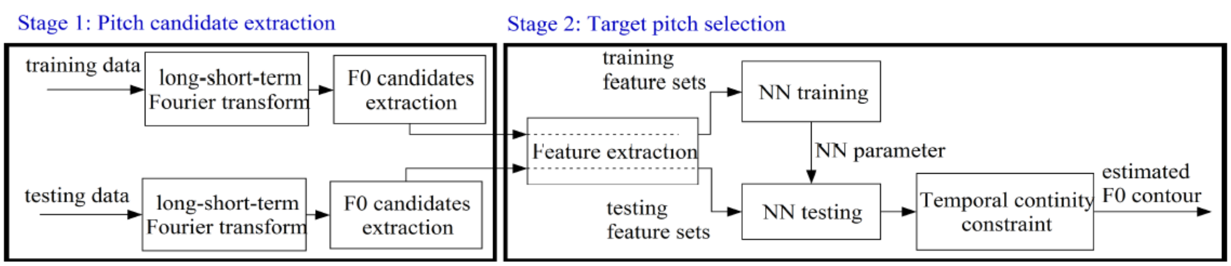


Figure 45: Block diagram of two-stage pitch estimation algorithm.

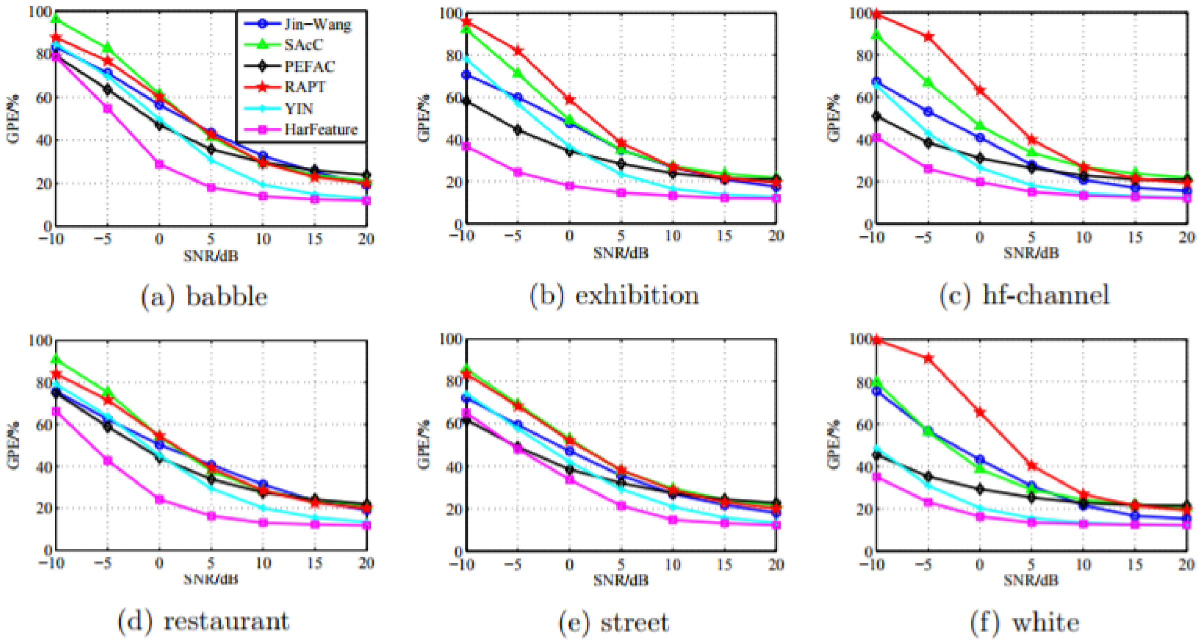


Figure 46: Gross Pitch Error (GPE) results for Keele database in different noise type at multiple SNR levels. These results demonstrate that the proposed HarFeature algorithm yielded substantially better performance (lower GPE) than the compared state-of-the-art algorithms across various types and levels of noise

Experimental evaluation of algorithms that improve speech intelligibility of CI listeners in noise: The essence of the developed algorithms is to estimate and preserve the inherent harmonic structures of speech signal, as well as use of harmonic structures as frequency domain cues to estimate the noise type and level. Traditional statistical based methods are usually effective in the stationary noise by tracking noise along the time dimension. The advantage of the proposed method is that it takes into account both voiced and unvoiced speech sections, and the computational complexity is also low. Thus the statistical based method MMSE was adopted as the framework for the overall speech enhancement. The developed speech enhancement algorithm was based on harmonic structure estimation combined with MMSE method for a leveraged solution. Harmonic structure was estimated in each frame for the target speech. Next, the noise estimation was performed based on exploring the harmonic structures of the target speech. Noise was estimated using the noise-dominated spectrum which is distributed between the adjacent harmonic partials. The estimated noise variance was then employed in the traditional MMSE framework for the *a-prior* and *posterior* SNR estimation to obtain the gain function for the target speech.

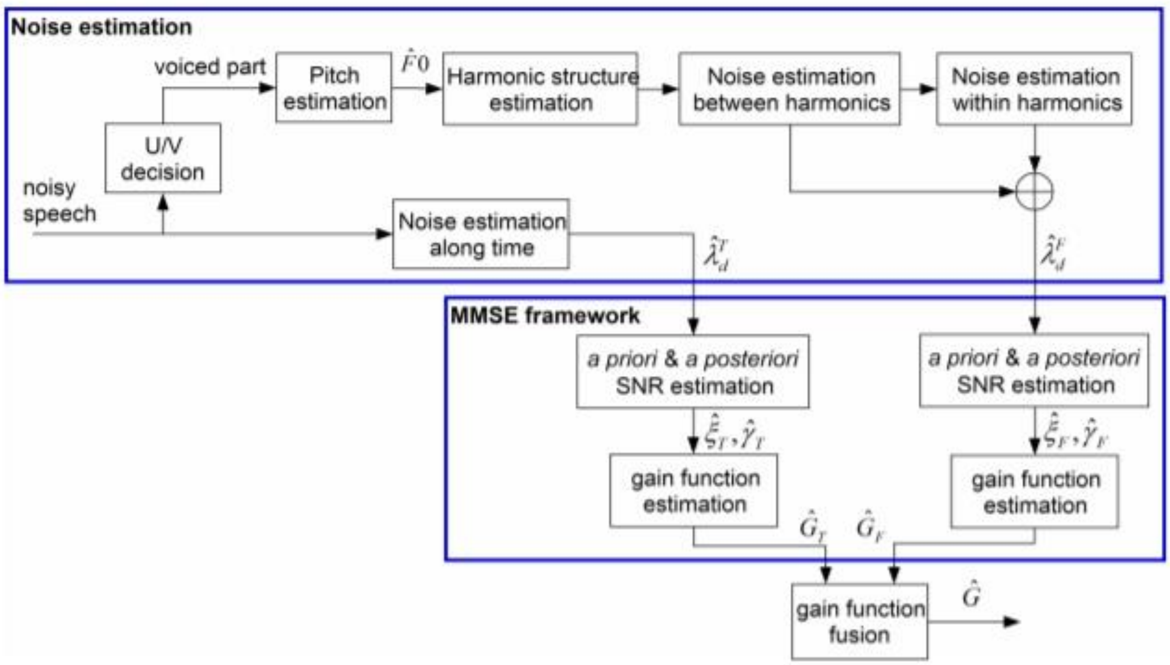


Figure 47: Block diagram of the speech enhancement algorithm

Listening experiments with CI subjects demonstrated the proposed algorithm was able to improve the speech intelligibility for both stationary and non-stationary noise.

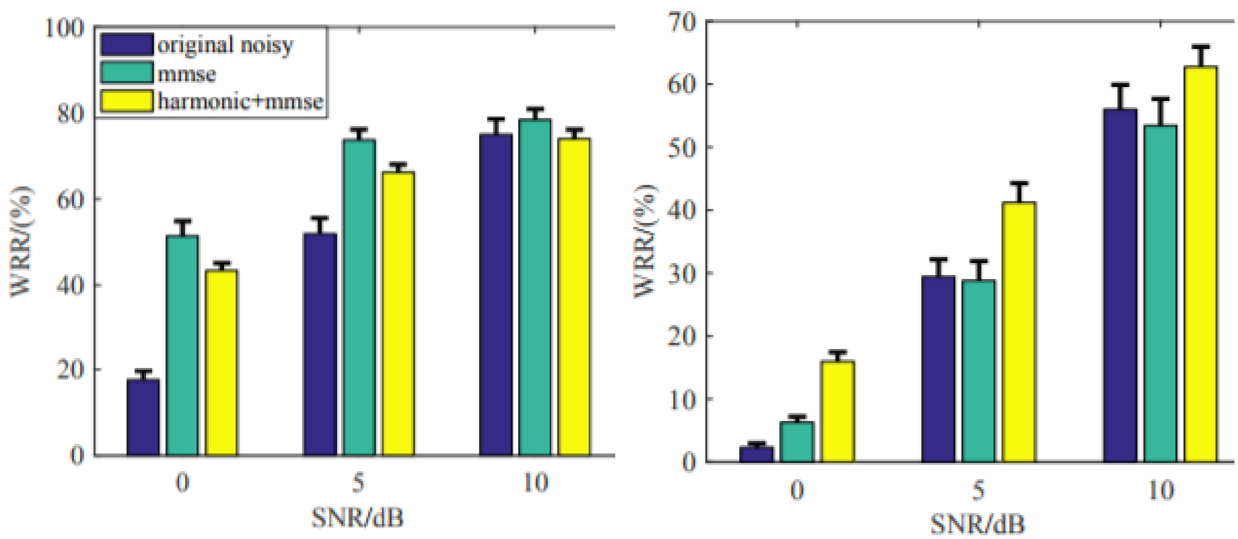


Figure 48: Word recognition rate with CI users (N=6), (a) in speech-shaped noise, (b) in babble noise.

3. Channel selection-based algorithms for noise suppression in cochlear implants.

The performance of current channel selection criteria used in cochlear implant devices (e.g., maximum selection criterion used in ACE) degrades significantly in the presence of noise. In noisy backgrounds, coding strategies that select the "right" channels for stimulation could potentially produce substantial improvements in intelligibility. A study was conducted to assess the performance of two alternative channel selection criteria in terms of intelligibility and subjective quality with CI users in noise. The performance was compared against that of the maximum selection scheme employed in the ACE strategy (comparison was also made with the CIS strategy). Sentences were presented to seven CI users in speech weighted noise (-5, 0, and 5 dB SNR). Both channel selection criteria were implemented under ideal conditions where a priori knowledge of the target and masker was assumed. This was done to assess the full potential benefit of selecting the "right" channels for stimulation in noisy backgrounds. Substantial intelligibility improvement relative to the CI users' daily strategy (i.e., ACE or CIS) was achieved with the two different channel selection criteria under all noisy conditions considered. No significant difference in subjective quality of noisy speech processed by the two channel selection criteria was observed.

Further details on the algorithm, methods, and evaluations can be found in: Oldooz Hazrati, and Philipos C. Loizou, “Comparison of Two Channel Selection Criteria for Noise Suppression in Cochlear Implants”, J. Acoust. Soc. Am., 133(3), 1615-1624, 2013.

4. Physiologically inspired algorithms for improved speech intelligibility with cochlear implant devices.

In daily communication, speakers aim to communicate their message in a manner that is intelligible to listeners. When individuals with normal hearing become aware of reduced auditory feedback due to environmental noise, they likely adopt a different speaking style called the ‘Lombard effect’. The Lombard effect is the tendency of speakers to modify their speech production while speaking in the presence of loud noise. Increased levels of masking noise lead to increase in vocal effort including energy, fundamental frequency, and glottal spectral slope. Lombard speech modification is aimed at providing the listener with in-creased speech intelligibility in challenging listening environments. The Lombard effect is also known to degrade automatic speech systems such as automatic speech recognition (ASR) and speaker identification (SID). While well-studied for normal hearing listeners and automatic speech systems, the Lombard effect has received little, if any, attention in the field of cochlear implant users.

To our knowledge, no study has examined whether cochlear implant users employ the Lombard effect during voice communication. This work provided a comprehensive investigation of the research concerning Lombard effect for cochlear implant users with post-lingual deafness. We investigated the nature of the Lombard effect that was produced by cochlear implant users. A variety of acoustic and phonetic features including voice power, fundamental frequency, glottal spectral tilt, phoneme duration, and formant frequencies were analyzed. Mobile personal audio recordings from continuous single-session audio streams collected over an individual’s daily life were used for these analyses. Prior advancements in this domain include the “Prof-Life-Log” longitudinal study at UT-Dallas (CRSS). The Lombard effect was observed in the speech production of all cochlear implant users. The results indicated that both suprasegmental (e.g., F0, glottal spectral slope and vocal intensity) and segmental (e.g., F1 for /i/ and /u/)) features were altered in such environments. Along with speech production characteristics, the research also focused on the effect of Lom-bard speaking style on intelligibility by cochlear implant users. A speech corpus for the perceptual experiments of Lombard speech was formulated with normal hearing speakers. Subjective listening tests were performed to provide how cochlear implant users respond to Lombard speech in challenging listening environments. The results indicated that the Lombard speech yielded a significant improvement in intelligibility in both quiet and noisy listening conditions. The specific modification of speech production of cochlear implant users under the Lombard effect may contribute to some degree an intelligible communication in adverse noisy environments. Lastly, a Lombard effect-inspired speech enhancement algorithm for cochlear implant users was developed. A previous proposed framework based on Source Generator theory was employed to perturb neutral speech production based on Lombard effect modification. Data from subjective evaluation demonstrated the effectiveness of the proposed speech enhancement algorithm. The results indicated improvement in intelligibility when providing neutral speech which was modified based on Lombard effect properties via the proposed algorithm. The specific variations due to Lombard effect can be leveraged for new algorithm development and further applications of speech technology to benefit cochlear implant users.

Development of a naturalistic speech corpus: The first step in order to investigate speech production of CI users in naturalistic environments was to develop a new speech corpus. Mobile personal audio recordings from continuous single-session audio streams were used to observe this effect over an individual's daily life. In the data collection, CI and NH speakers produced read and conversational speech over various UT-Dallas on-campus environments (e.g., office and cafeteria). Following the controlled locations, additional audio recordings were collected in uncontrolled environments (outside campus, e.g., home and restaurants). A set of acoustic and orthographic transcription labels were assigned by a human transcriber. Phoneme-level transcription labels were assigned automatically by forced alignment procedures. A total of 100 hours of personal audio recordings were collected from 6 cochlear implant and 18 normal hearing participants. This collection of speech production provided a unique and unprecedented opportunity to explore real-world listening and speech in diverse environments for CI-to-NH communications.

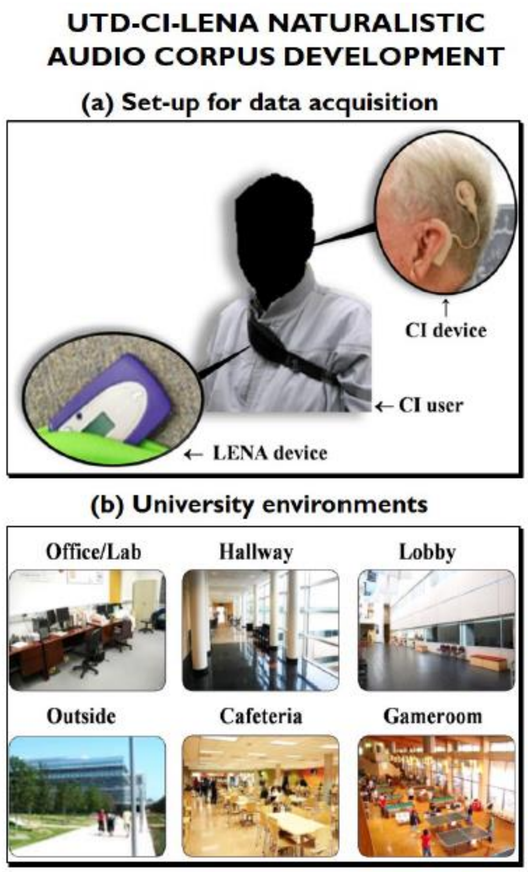


Figure 49: Naturalistic data collection for cochlear implant subjects: (a) set-up for data acquisition using the LENA unit, and (b) naturalistic environments on UT-Dallas college campus for data collection.

Effect of environmental noise on speech production of cochlear implant users – a naturalistic study: The second dissertation contribution was that it analyzed speech production of cochlear implant adults with respect to environmental changes. The analysis conducted in this work were based on mobile personal audio recordings collected over various realistic environments on university campus. The parametric variations in vowel, consonant and individual phoneme production were investigated as a function of varying environments. Data from these analyses indicated that Lombard effect was found in speech of post-lingually deafened cochlear implant adults. Speakers demonstrated increased vocal effort, including F0 and speech power, as well as altered glottal spectral slope, and phoneme duration in response to challenging noisy environments. Segmental articulatory movements, for example, F1 for specific phonemes such as /a/, /æ/, /i/, and /u/, also appeared to play an important part in speech production under noise. The significance of the results is that the Lombard effect could potentially help cochlear implant users to ensure/maintain intelligible communication by compensating for the reduced signal-to-noise ratio. Lombard speech research in this research may shed light on the underlying mechanisms of speech production of cochlear implant users.

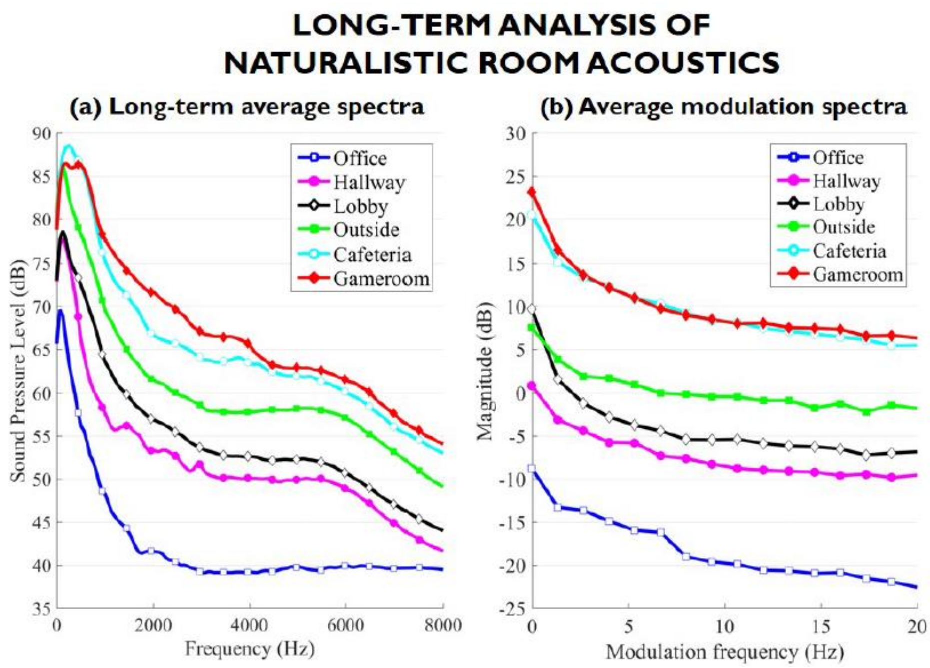


***Figure 50: Long-term analysis for six naturalistic environments used in this study on UT-Dallas campus:** (a) long-term average spectra, and (b) average modulation spectra. Each line corresponds to the average long-term features for each naturalistic environments. The main noise sources were recorded prior to subject's speech production for 3 minutes in each environment.*

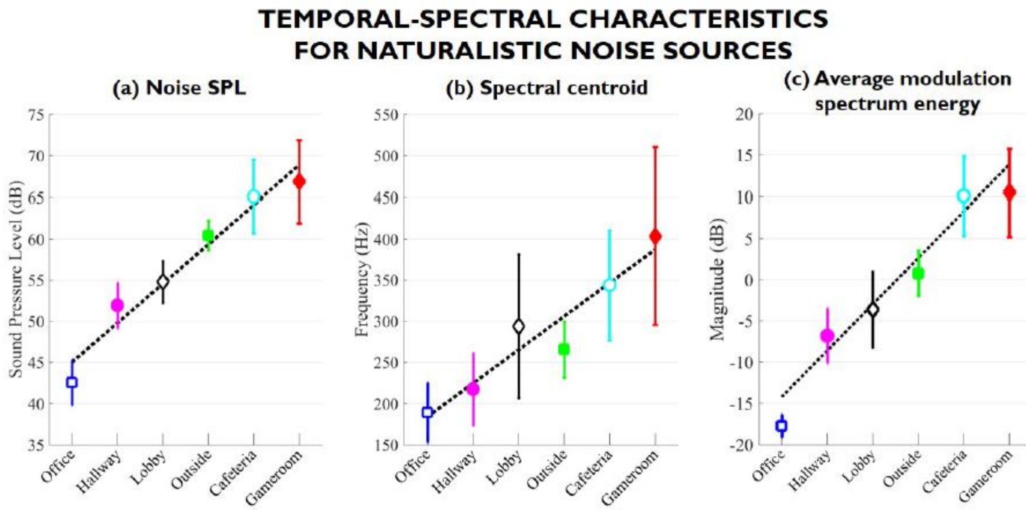


***Figure 51: Acoustic characteristics of background noises:** average (a) noise SPL, (b) spectral centroid, and (c) average modulation spectrum energy with respect to different environments. While average SPL shows changes in signal strength over time-domain, spectral centroid represents where spectral energy was concentrated in frequency-domain. Average modulation spectrum energy estimate the relative degree of stationarity for the noise signal.*

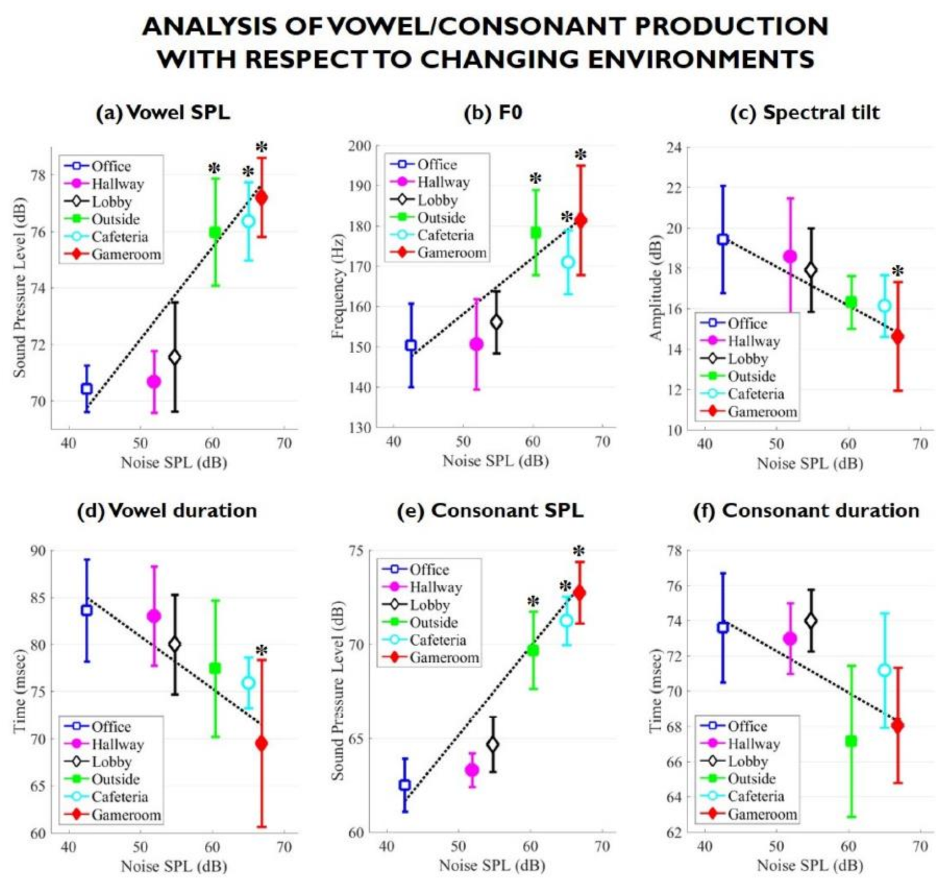


***Figure 52: Acoustic analysis of vowel and consonant productions:** individual variations of (a) vowel SPL, (b) fundamental frequency (F0), (c)spectral tilt, (d) vowel duration, (e) consonant SPL, and (f) consonant duration, as a function of varying noise SPL. Asterisk indicates statistical significance ($p< 0.05$) from neutral speech.*

Influence of Lombard Effect on speech intelligibility of cochlear implant users: Next, the perceptual analysis of Lombard effect by post-lingually deafened cochlear implant users was conducted. We analyzed acoustic properties of Lombard and neutral speech from two normal hearing speakers. A speech corpus that is intended for the perceptual experiments of Lombard speech was developed for this analysis. Subjective intelligibility was measured with cochlear implant users across signal-to-noise ratios. Data from the perceptual analysis suggested Lombard speech had a higher intelligibility than neutral speech for electric hearing in both quiet and noisy environments. An advantage of +7.6 to +13.2 percentage points in intelligibility for Lombard speech over neutral speech. Larger improvement in intelligibility was found in challenging listening environments. Speech with higher vocal effort were more intelligible than the speech produced with lower vocal effort. These results suggested that the modification of speech production parameters obtained un-der noise contribute to the Lombard speech intelligibility. The findings presented in this study further our understanding of the speech perception of cochlear implant users.

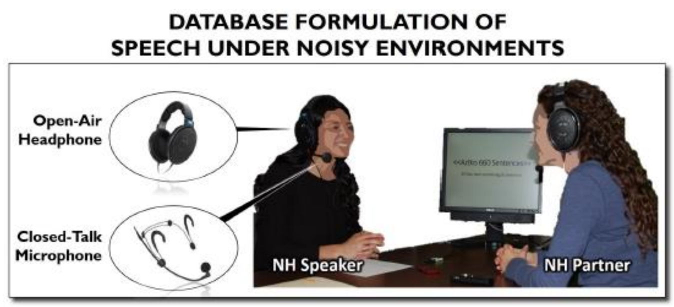


Figure 53: Lombard effect data collection with normal hearing participants. *Data collection was performed in a sound recording booth. A set-up for data acquisition using open-air headphone and close-talk microphone is demonstrated.*

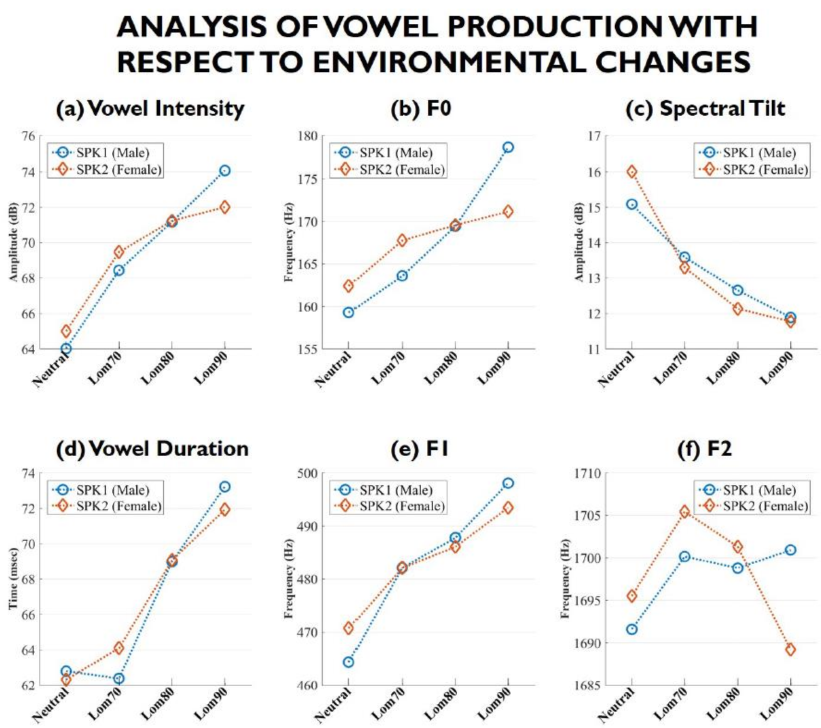


Figure 54: Acoustic analysis of vowel production*: individual variations of (a) vowel intensity, (b) fundamental frequency (F0), (c) spectral tilt, (d) vowel duration, (e) first formant (F1), and (f) second formant frequency (F2), as a function of varying listening environments. Asterisk indicates statistical significance ($p < 0.05$) from natural speech.*

Development of Lombard perturbation strategy for improving speech intelligibility of cochlear implant users: The above analyses helped shed light on Lombard characteristics of everyday speech. As a next step, a Lombard effect-based speech enhancement algorithm for cochlear implant users was developed. The proposed algorithm controlled the acoustic parameters of neutral speech to present Lombard speaking style to improve intelligibility in noisy environments. Acoustic variations for neutral and Lombard conditions were modeled and used to generate Lombard synthetic speech. The modification areas considered in the proposed algorithm were voice intensity, overall spectral contour, and sentence duration. The analysis and modeling conducted here was based on a previously established framework, Source Generator theory. Subjective listening evaluation was performed with five cochlear implant users to demonstrate the effectiveness of the speech modification algorithm. The results from perceptual analysis indicate potential of the Lombard effect based speech modification algorithm with perceptual benefits to cochlear implant users in noisy environments. The specific knowledge provided in this study can have practical implications for developing speech enhancement algorithms for cochlear implant users.

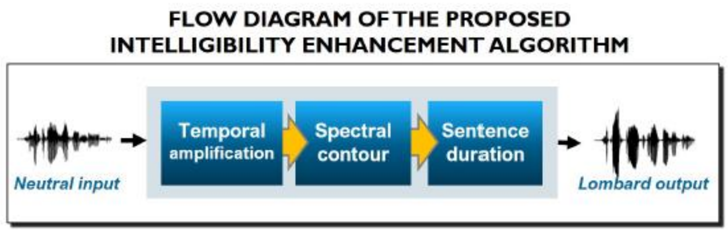


Figure 55: Block diagram of the developed speech enhancement algorithm. The algorithm controls the acoustics features of input neutral speech to present Lombard speaking style output. The modification areas considered here were: (1) amplification, (2) spectral contour, and (3) overall sentence duration.

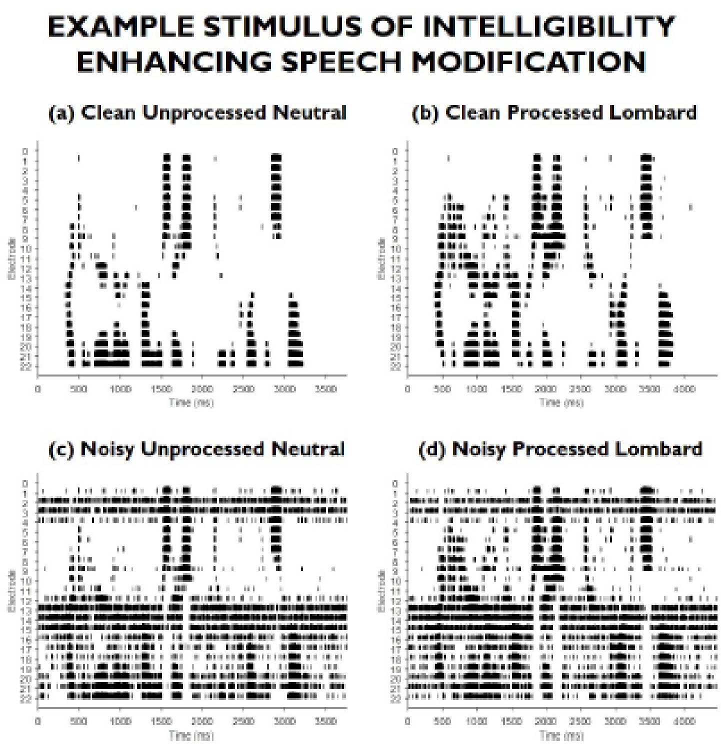


Figure 56: Stimulus output patterns (electrodograms) of the sentence "Basketball can be an entertaining sport": (a) original neutral speech reference, (b) Lombard processed speech via the proposed algorithm, (c) neutral speech mixed with speech-shaped noise at 10 dB SNR, and (d) Lombard processed speech mixed with speech-shaped noise at 10 dB SNR.

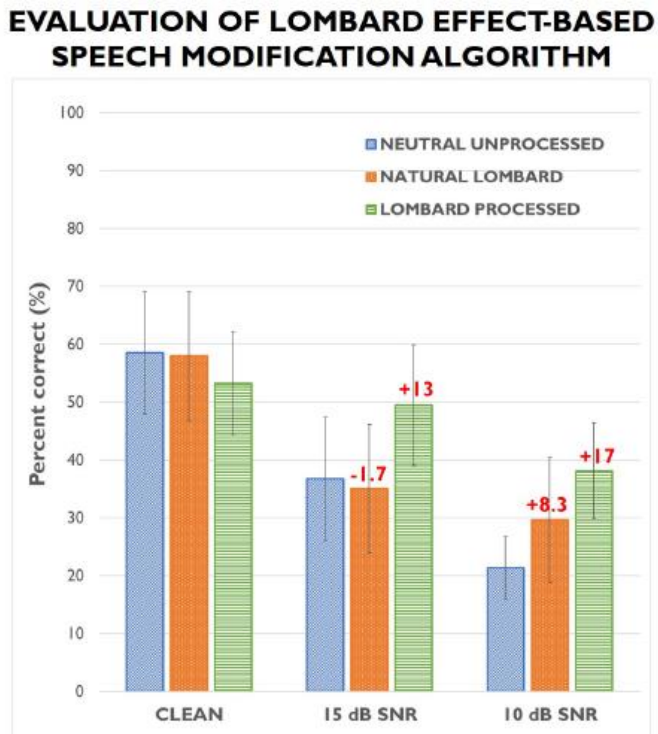


***Figure 57: Word intelligibility scores for unprocessed neutral, natural Lombard, and processed Lombard speech with five cochlear implant users as a function of signal-to-noise ratio.** The Lombard processed speech was generated by modifying the neutral speech via the proposed intelligibility-enhancing algorithm. The neutral unprocessed and natural Lombard speech were obtained by normal-hearing subjects while speaking in quiet and in large-crowd noise at 90 dB SPL respectively.*

[Customized sound processing and fitting strategies for cochlear implants]

Although electrode placements relative to the spiral ganglion are generally unknown, conventional sound coding algorithms and clinical fitting procedures follow a one-size-fits-all strategy. In the scope of this project the aim was to customize/personalize the sound coding and fitting procedures according to an individual's cochlear physiology. This was achieved by utilizing novel imaging procedures, more specifically CT images of recipients' cochleae, to determine the precise spatial location and orientation of the electrode contacts and the corresponding neural stimulation sites to produce a tailor-fit frequency-to-place function. In addition, patient-specific channel selection optimization techniques have been presented that aim to optimize presentation of electrical stimulation patterns. The proposed schemes were evaluated in groups of normal hearing individuals using CI simulations as well as cochlear implant recipients in collaboration with Vanderbilt University Medical Center. The data from the experiments suggest that patient-specific sound coding and fitting schemes may potentially aid in achieving higher asymptotic performance and possibly faster adaptation to electric hearing.

In a cochlear implant, any range of frequency can theoretically be presented to any electrode. 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. 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. This frequency mismatch (between intended and actual pitch perception) holds potential for reducing the efficacy of coded speech information to the auditory cortex and, consequently, limits speech recognition. In this work, we proposed a patient-specific frequency assignment strategy which helps to minimize sub-optimal frequency-place mapping distortions in CIs. The algorithm leverages image-guided procedures to determine the true location of individual electrodes with respect to the nerve fibers and tailor-fits a frequency place function based on an individuals' electrode-neuro interface.

The proposed strategy developed was shown to utilize pre and post implantation CT scans of the recipients' cochleae to determine the precise spatial location of electrode contacts and their corresponding neural stimulation sites and thus generate an optimal user-customized frequency-place function which is used to derive frequency characteristics of the filter banks. This was achieved by maximizing the frequency match at lower frequencies (frequency range of first three formants), and introduced a mild compression as needed to avoid truncation (e.g., due to shallow insertion). Mid and high frequency bands were assigned along a conventional logarithmic filter spacing.

The proposed custom frequency assignment scheme was evaluated in the following three listening studies:

i) Study 1: Acoustic simulations of cochlear implants with normal hearing listeners:

Performance of the proposed strategy was evaluated with 42 normal hearing (NH) listeners using acute acoustic simulations of a cochlear implant with actual electro-neuro image maps of CI users. The simulation data indicated significantly better speech recognitions scores as compared to the default clinical mapping scheme on all measures of speech. Although the improvements were observed for all image maps that had various degrees of frequency mismatch, the proposed strategy produced significant improvements in moderate-to-extreme frequency-place mismatch conditions.

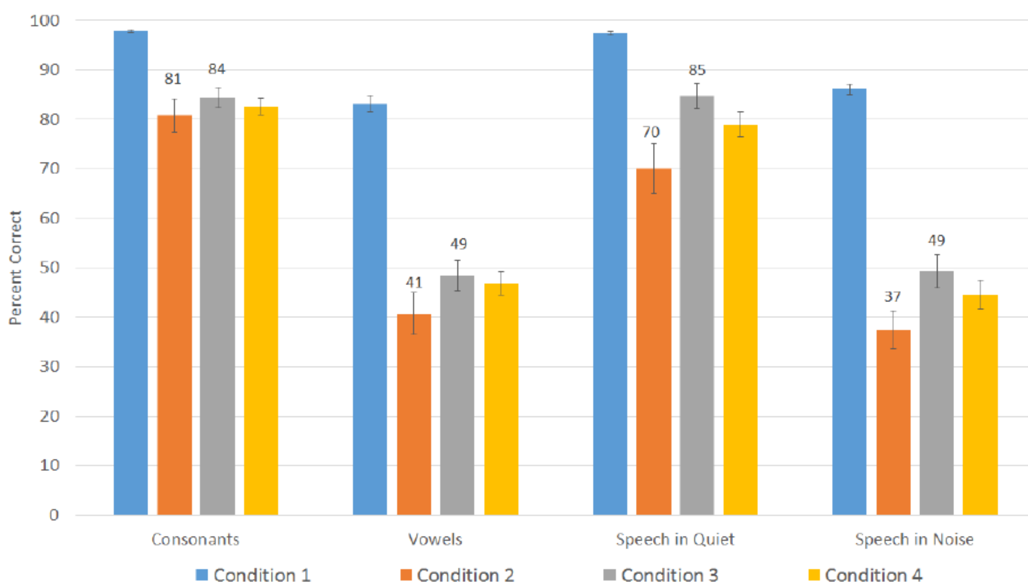


Figure 58: Average speech recognition scores of 42 participants from Study 1 with normal hearing individuals. Percentage correct scores for recognition of consonants, vowels, speech in quiet, and speech in noise (SNR = 10 dB) with respect to four frequency mapping conditions. Condition 1: Default Frequency allocation, with ideal electrode positioning. Condition 2: Default Frequency allocation, with true electrode positioning. Condition 3: Custom frequency allocation, with true electrode positioning. Condition 4. Frequency allocation matched with true electrode positioning. Error bars represent standard errors of means.

ii) Study 2: Simulating the effect of adaptation to frequency-place functions in normal hearing listeners:

Since acute simulation scores may underestimate the true potential effects of learning and user-adaptation to speech with a reduced and degraded set of spectral cues, the proposed technique was also investigated in a semi-chronic paradigm. Ten normal-hearing listeners were provided with approximately four hours of auditory training and tested at different intervals with both the default frequency mapping scheme and the proposed custom mapping strategy. On all measures of speech, all participants showed significantly better performance with the proposed custom mapping strategy, at least within the four hour test period. The data indicated that listeners adapt to both schemes, but the level of accommodation and final level of performance with the proposed scheme was still significantly better than the default clinical strategy.

The above two simulation studies with normal hearing subjects served as a viable proof-of-concept for the follow-on investigation with cochlear implant users.

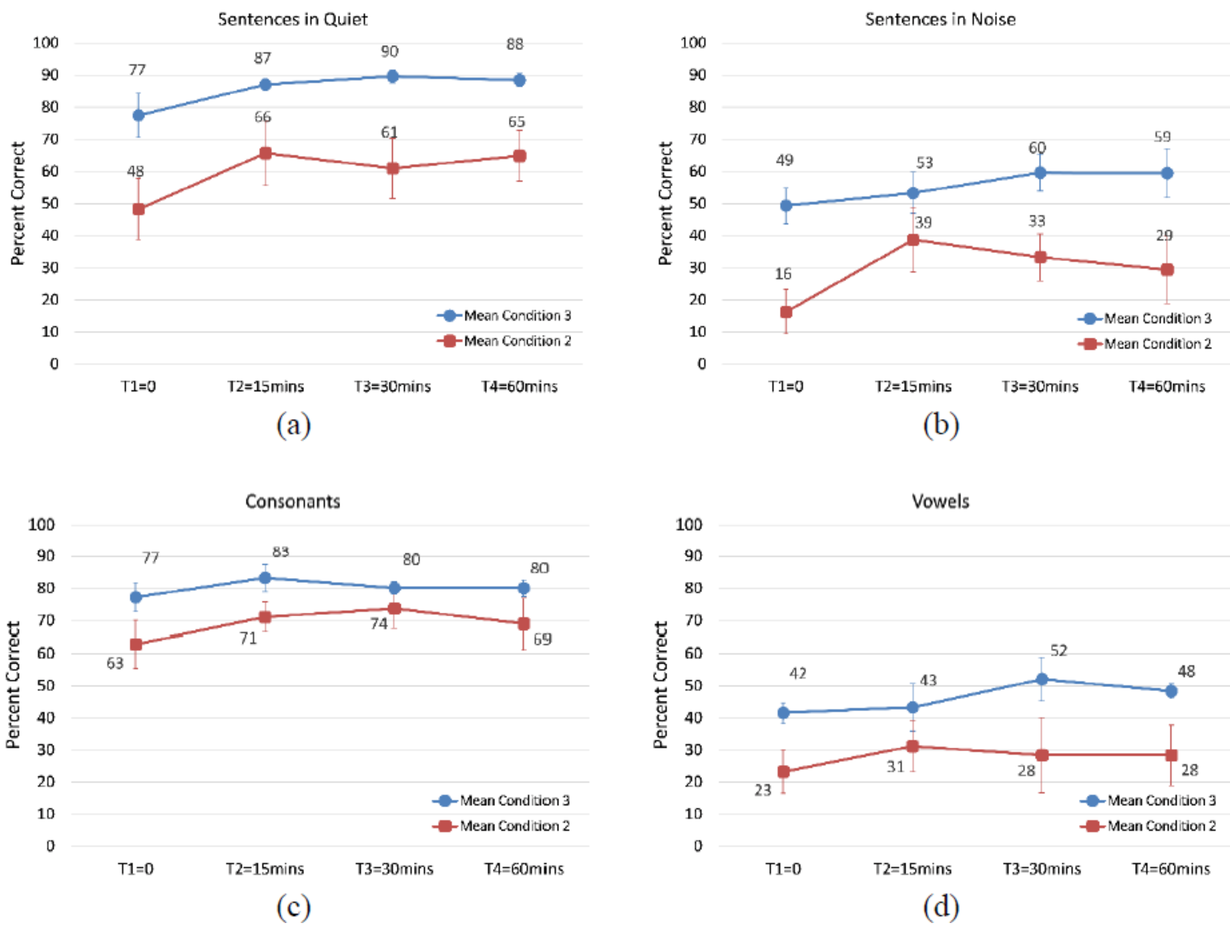


Figure 59: Average test scores measured at four time intervals in study 2. Subjects were tested acutely (with minimal training) at the start of the test, and then progressively given audio/video (A/V) training sessions and tested at intervals 2, 3, and finally at 4. Condition 2 mimics default frequency allocation with true tonotopic place, Condition 3 depicts proposed custom frequency allocation scheme with true tonotopic place.

iii) Study 3: Evaluation with cochlear implant recipients:

Five experienced, post-lingually deafened adult CI users participated in this semi-chronic study. Imaging data of the recipients' cochleae provided the needed relationships between spatial location of electrode contacts and the characteristic frequencies of the nerve fibers. These physiological blueprints were used to optimize and "tailor-fit" frequency-place functions for each participant.

Each participant's performance was measured on ten speech recognition tasks with their clinically assigned frequency map, as well as with the proposed custom frequency assignment scheme. Participants used the experimental program for three months and their performance levels were measured at three stages during the study, i) acutely, ii) after 1 week, and iii) after 3 months. Consistent with typical clinical observations, performance levels with the experimental program dropped significantly lower than the clinical processor for all CI users. However, all subjects showed progressive improvement with extended use of the experimental programs on all measures of speech recognition. The performance levels improved after 1-week post-activation, and were not significantly different than their baseline scores with the clinical maps. At three months, performance levels with the experimental programs were significantly better than the acute scores and were also not statistically different than their original clinical strategy. Overall the progressive improvement with the experimental programs indicated the effects of learning and at least partial adaptation to the custom frequency maps. By the end of the study, all participants chose to keep the experimental program either exclusively, or with their old clinical programs.

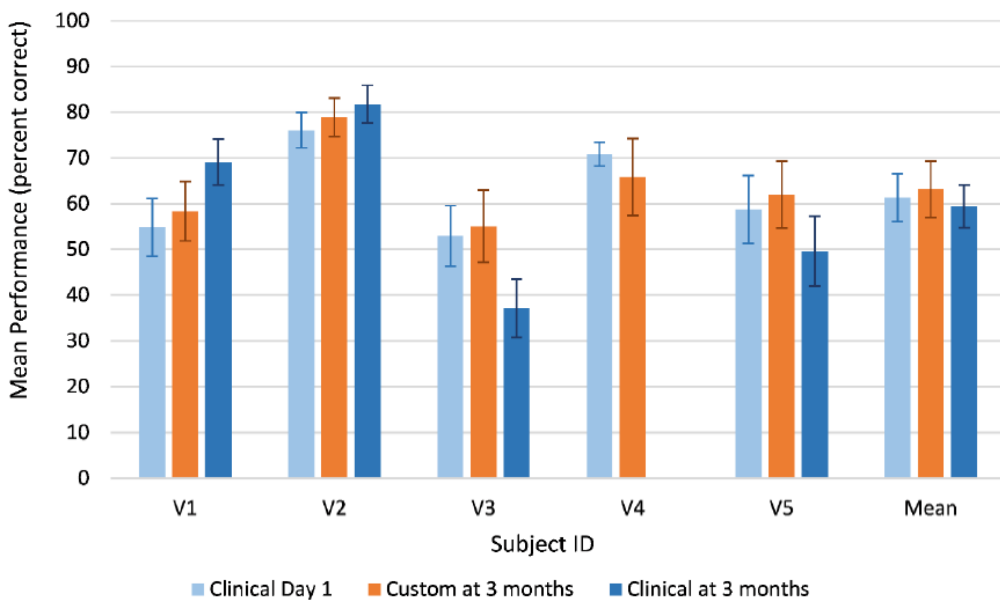


Figure 60: Mean performance of 5 CI subjects on 10 speech recognition tasks with clinical and custom frequency maps. Error bars represent standard errors of means

[Speech recognition with cochlear implants in natural listening environments]

A number of studies were preformed to assess speech intelligibility with cochlear implant devices in naturalistic acoustic environments. Some of these peer-reviewed studies are given as follows:

Evaluation and analysis of whispered speech for cochlear implant users: Gender identification and intelligibility: This study investigated the degree to which whispered speech impacts speech perception and gender identification in cochlear implant (CI) users. Listening experiments with six CI subjects under neutral and whispered speech conditions using sentences from the UT-Vocal Effort II corpus (recordings from male and female speakers) were conducted. Results indicated a significant effect of whispering on gender identification and speech intelligibility scores. In addition, no significant effect of talker gender on the speech/gender identification scores was observed. Results also suggested that exposure to longer speech stimuli, and consequently more temporal cues, would not improve gender identification performance in CI subjects.

Further details on the algorithm, methods, and evaluations can be found in: Oldooz Hazrati, Hussnain Ali, John H.L. Hansen, and Emily Tobey, "Evaluation and Analysis of Whispered Speech for Cochlear Implant Users: Gender Identification and Intelligibility", J. Acoust. Soc. America, vol. 138, No. 1, 74-79, July 2015.

Predicting the speech reception threshold of cochlear implant listeners using an envelope-correlation based measure: A modulation-based index was proposed for predicting speech intelligibility by cochlear implant (CI) listeners. The input to the proposed index were speech envelopes extracted using the individual CI user's daily strategy, and as such, this approach incorporates information about the number of active electrodes, shape of the compression function and electrical dynamic range. High correlation (r = 0.96) was achieved with the proposed index when evaluated with speech-reception thresholds (SRTs) obtained by CI users in steady and speech-masker conditions. This outcome suggests that the information contained in electrodograms seems to be sufficient for reliably predicting CI user's performance in noise. The proposed index can be used by clinicians to optimize the selection of fitting parameters of individual CI users for better performance in noise.

Further details on the algorithm, methods, and evaluations can be found in: Nima Yousefian, Philipos C. Loizou, "Predicting the speech reception threshold of cochlear implant listeners using an envelope-correlation based measure," The Journal of the Acoustical Society of America 2012 132:5, 3399-3405.

Evaluation of adaptive dynamic range optimization in adverse listening conditions for cochlear implants: The aim of this study was to investigate the effect of Adaptive Dynamic Range Optimization (ADRO) on speech identification for cochlear implant (CI) users in adverse listening conditions. In this study, anechoic quiet, noisy, reverberant, noisy reverberant, and reverberant noisy conditions were evaluated. Two scenarios were considered when modeling the combined effects of reverberation and noise: (a) noise is added to the reverberant speech, and (b) noisy speech is reverberated. CI users were tested in different listening environments using IEEE sentences presented at 65 dB sound pressure level. No significant effect of ADRO processing on speech intelligibility was observed.

Further details on the algorithm, methods, and evaluations can be found in: Hussnain Ali, Oldooz Hazrati, Emily Tobey, and John H.L. Hansen, "Evaluation of Adaptive Dynamic Range Optimization in Adverse Listening Conditions for Cochlear Implants", J. Acoust. Soc. Am., 136(3), EL242-EL248, 2014.

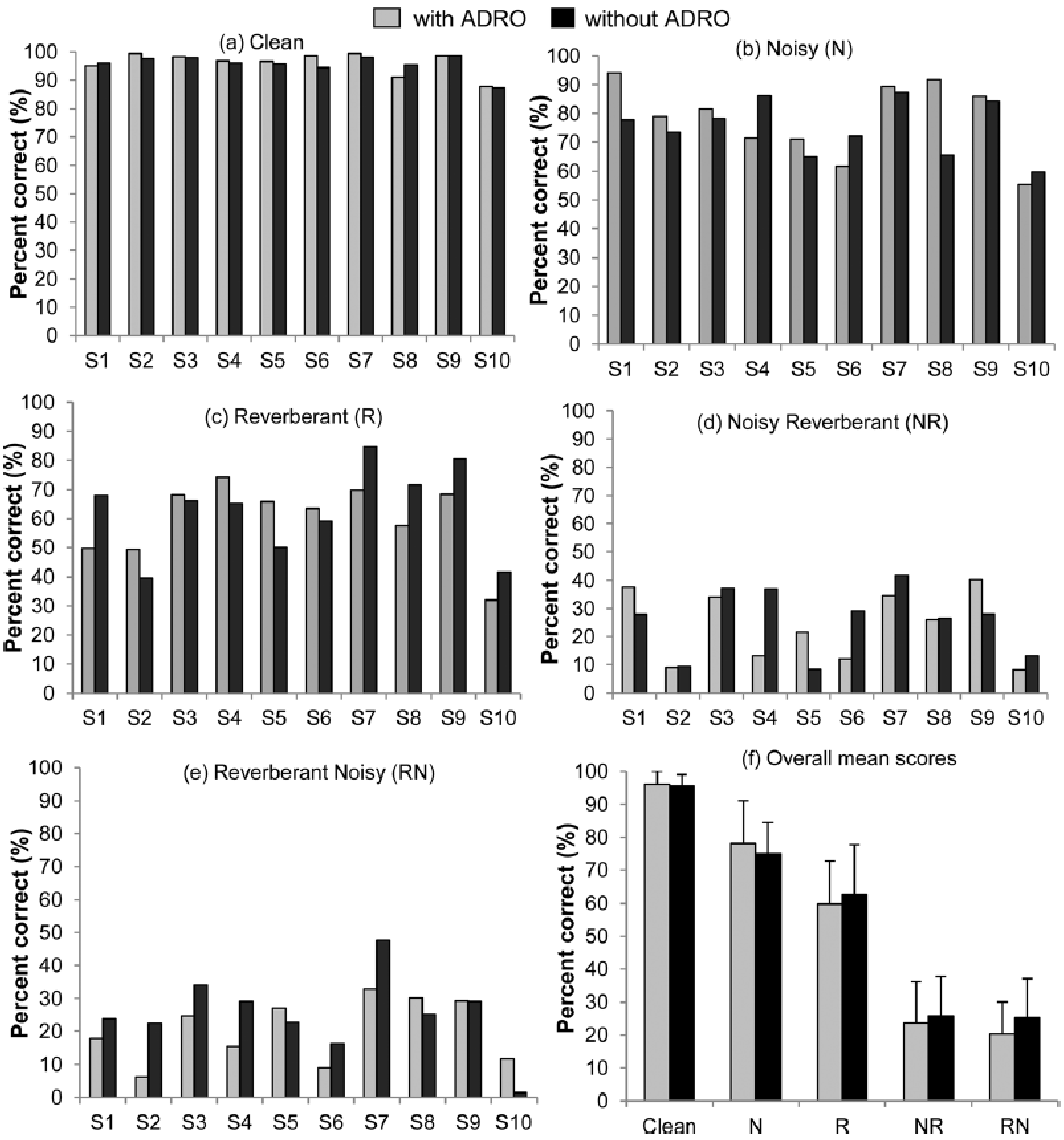


Figure 61: Individual speech intelligibility scores of ten CI users in (a) anechoic quiet (clean), (b) noisy (N, SNR¼10 dB), (c) reverberant (R, T60¼0.6 s), (d) noisy reverberant (NR, T60¼0.6 s, RSNR¼10 dB), and (e) reverberant noisy (RN, SNR¼10 dB, T60¼0.6 s) conditions. Panel (f) demonstrates average scores in all conditions. The error bars in panel (f) indicate standard deviations.

Objective speech intelligibility measurement for cochlear implant users in complex listening environments: Objective intelligibility measurement allows for reliable, low-cost, and repeatable assessment of innovative speech processing technologies, thus dispensing costly and time-consuming subjective tests. To date, existing objective measures have focused on normal hearing model, and limited use has been found for restorative hearing instruments such as cochlear implants (CIs). In this work, we evaluated the performance of five existing objective measures, as well as proposed two refinements to one particular measure to better emulate CI hearing, under complex listening conditions involving noise-only, reverberation-only, and noise-plus-reverberation. Performance was assessed against subjectively rated data. Experimental results showed that the proposed CI-inspired objective measures outperformed all existing measures; gains by as much as 22% could be achieved in rank correlation.

Further details on the algorithm, methods, and evaluations can be found in: Joao F. Santos, Stefano Cosentino, Oldooz Hazrati, Philipos C. Loizou, and Tiago H. Falk, "Objective Speech Intelligibility Measurement for Cochlear Implant Users in Complex Listening Environments", Speech Commun., 55(7), 815-824, 2013.

The combined effects of reverberation and noise on speech intelligibility by cochlear implant listeners: The purpose of this study was to assess the individual effect of reverberation and noise, as well as their combined effect, on speech intelligibility by CI users. Sentence stimuli corrupted by reverberation, noise, and reverberation + noise were presented to 11 CI listeners for word identification. Subjects were tested in two reverberation conditions (T60 = 0.6 s, 0.8 s), two noise conditions (SNR = 5 dB, 10 dB), and four reverberation + noise conditions. Results indicated that reverberation degrades speech intelligibility to a greater extent than additive noise (speech-shaped noise), at least for the SNR levels tested. The combined effects were greater than those introduced by either reverberation or noise alone. The conclusions of this study were as follows: The effect of reverberation on speech intelligibility by CI users was found to be larger than that by noise. The results from this study highlighted the importance of testing CI users in reverberant conditions, since testing in noise-alone conditions might underestimate the difficulties they experience in their daily lives where reverberation and noise often coexist.

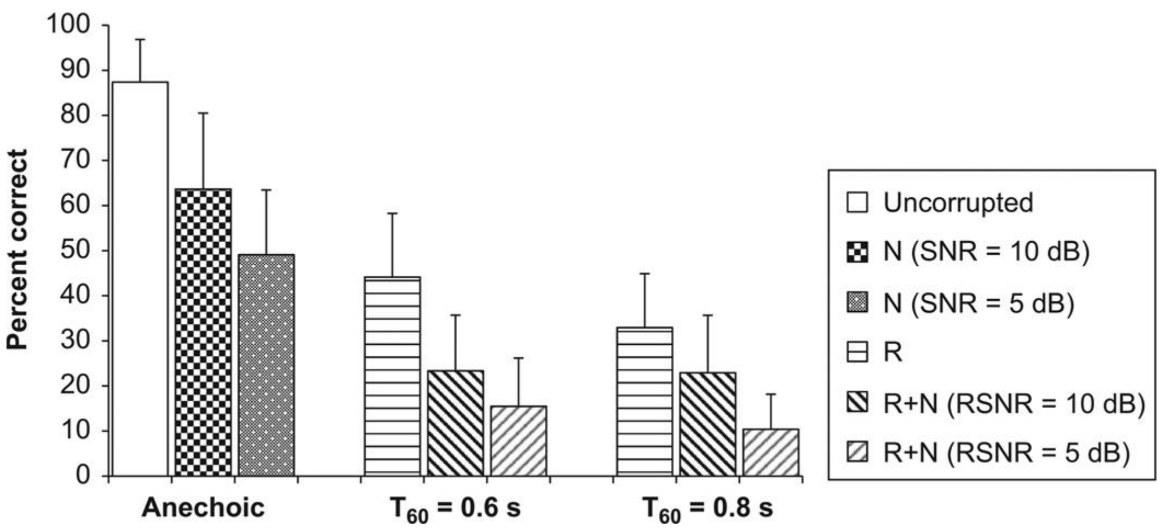


Figure 62: Average percent correct scores of CI users (n = 11) as a function of reverberation time and SNR/RSNR level. Error bars indicate standard deviations. Results indicate detrimental effects of noise and reverberation on speech recognition with CI devices.

Further details on the algorithm, methods, and evaluations can be found in: Oldooz Hazrati, and Philipos C. Loizou, "The Combined Effect of Reverberation and Noise on Speech Intelligibility by Cochlear Implant Listeners", Int. J. Audiol., 51(6), 437-443, 2012.

Chimeric experiments to investigate the perceptual importance of temporal envelop and temporal fine structure between tonal and non-tonal language: Speech intelligibility is more correlated with temporal envelope than temporal fine structure in non-tonal language. In cochlear implant devices, the encoding technique is based on temporal envelope rather than temporal fine structure. However, for tonal-language, the pitch pattern information is related with the specific tone information. In this work, we compared the perceptual difference of temporal envelope and temporal fine structure between tonal language and non-tonal language. The general trends of the experimental results for each language were similar between American English and Mandarin Chinese. There was no difference between the two languages for TE and TFS perceptual importance when the interfering signal was speech-shaped noise. However there was a significant difference for the perceptual importance of TE and TFS between the tonal language and non-tonal language when the interfering signal was competing speech. The results showed that compared to a non-tonal language, the tonal language is more sensitive to temporal fine structure distortion since the temporal fine structure contains tone information. The results of this study present guidelines for the CI encoding techniques for tonal-languages to obtain better speech intelligibility.

Effect of consonant-vowel boundary to speech perception in cochlear implants: The brain can perceptually restore inaudible portions of speech by using syntactic, semantic and linguistic knowledge. Such top-down repair enhances speech intelligibility in noisy environments. In this study, we explored the effect of phonemic restoration by presenting speech stimuli to CI users replaced with silence and noise utterances in consonant and vowel regions. The aim of the study was to investigate the restoration benefit as a function of speech degradation in CI users. The restoration dependency on the type of masked phoneme and type/duration of masker are particularly investigated. The contribution of consonants and vowels to speech perception in cochlear implant (CI) listeners is not well understood. A segment replacement paradigm on TIMIT speech corpus was used to evaluate the effect of consonant-vowel boundary to speech perception in CI listeners. Two processing strategies were created to emphasize the duration of vowels and consonants by presenting different amounts (0%, 40%, 60%, 80%, and 90%) of consonants and vowels. The first strategy (FVXC) preserved full vowels (FV) and presented different amounts of consonant (FV+%C) by replacing the consonant centers with either silence or speech shaped noise. The second strategy (FCXV) preserved full consonants (FC) and presented different amounts of vowel (FC+%V) by replacing the vowel centers with either silence or speech shaped noise. Clean speech and interrupted speech (3 Hz interruption rate, 50% duty cycle) were also presented. A PDA based research platform was used to present the stimuli to CI listeners.
For FCXV presentation condition, noise replacement paradigm had higher speech intelligibility compared to silence for 0%, 40%, and 60% vowel presentation conditions. There was no difference in speech intelligibility for 80% and 90% vowel presentations. Moreover, there was no significant difference between the replacement paradigms for FVXC presentation condition. Also, the speech intelligibility was very low (< 10%) for 3Hz interruption suggesting the inability of the CI listeners to tolerate interruptions in continuous speech. Overall it was concluded that the CI listeners have difficulty in fusing interrupted speech signals into one coherent speech stream and phonemic restoration in CI listeners occurs depending on the presented speech stimuli.

Further details on the algorithm, methods, and evaluations can be found in: Nirmal Srinivasan, Hussnain Ali, Emily Tobey, and Philipos C. Loizou, “Effect of consonant-vowel boundary to speech perception in cochlear implants,” Conference on Implantable Auditory Prostheses'13, (CIAP 2013) Lake Tahoe, California, USA, July 2013.

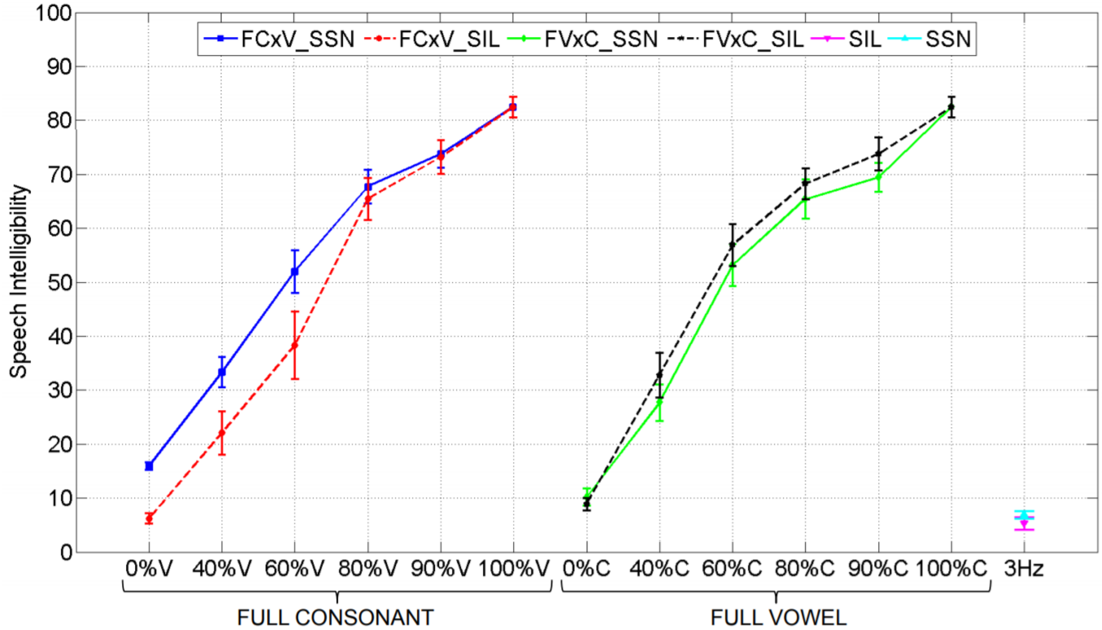


***Figure 63: Mean speech intelligibility for all experimental conditions.** The original TIMIT C-V boundary is at 0%V and 0%C. Error bars display standard error of mean.*

Amplitude modulation detection with cochlear implants: Time varying envelopes of the stimuli are key perception cues for both acoustic and electric hearing. CI users take advantage of their electric dynamic range to sense these amplitude modulations for sound perception. Most of the research in amplitude modulation detection (AM) with CI users has been focused on modulation frequencies greater than 50 Hz. A study was conducted to investigate the AM detection capability of CI users at low modulation frequencies (2Hz - 64 Hz) at different dynamic ranges (DR) (50% and 70%). 10 CI subjects were tested with AM detection experiment. Figure shows mean modulation detection thresholds of 10 subjects as a function of modulation frequency at 50% and 70% DR.

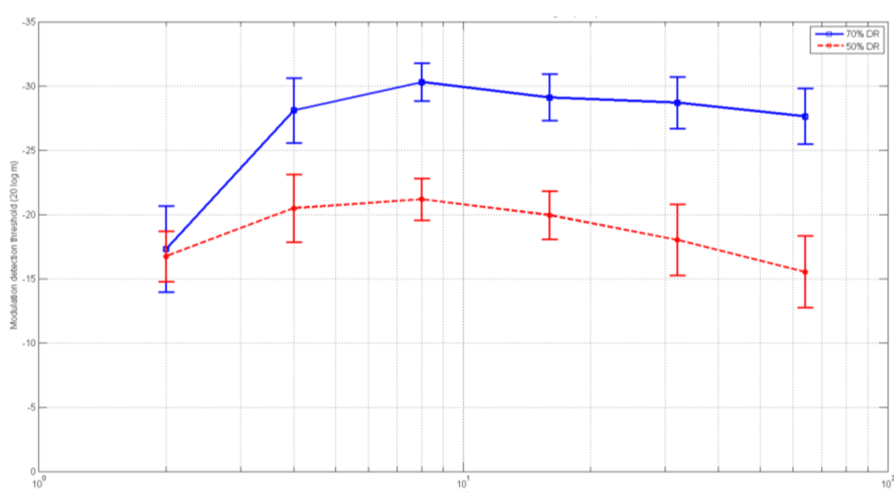


***Figure 64: Modulation detection thresholds as a function of modulation frequency** at 50% DR (shown in red) and 70% DR (shown in blue). Error bars represent standard error of the mean (SEM)*

[Bilateral Cochlear Implants]

Bilateral cochlear implants (BiCIs) provide significant spatial hearing benefits over a single implant, such as improved sound localization and speech-in-noise recognition. However, BiCI users still exhibit spatial hearing difficulties compared to normal hearing listeners. Deficits have been posited to result from the fact that the two implants function independently and that the lack of coordination between devices disrupts the ability to encode important acoustical cues used for spatial hearing. The UT-Dallas ciPDA device has a unique ability to provide a real-time synchronized stimulation across the ears, i.e., one time-clock is used to stimulate both implants. Experiments were preformed to assess the potential benefits of synchronized bilateral electrical stimulation and its impact on speech intelligibility of cochlear implant users.

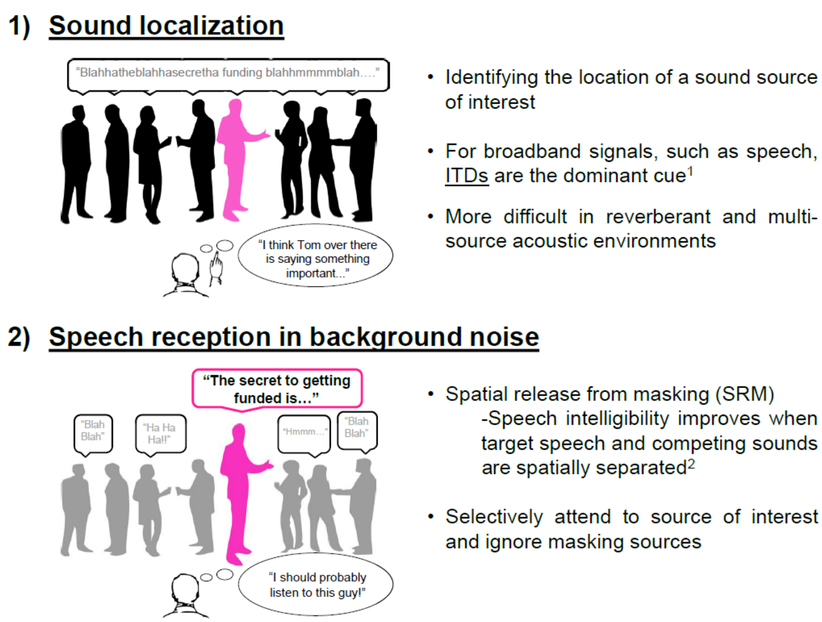


Figure 65: Sound localization and speech reception with bilateral cochlear implant devices.

Spatial hearing abilities in BiCI users listening with the ciPDA device: Litovsky and colleagues evaluated free-field sound localization and spatial release from masking (SRM; speech understanding when competing speech and target speech are co-located vs. spatially separated) using the ciPDA device. Testing was conducted in 10 post-lingually deafened BiCI users fitted with Cochlear Freedom or N5 devices and had a minimum of 1 year bilateral experience. Localization performance was measured by calculating root-mean-square (RMS) errors between target and response angles. Speech reception thresholds (SRTs) were measured for two loudspeaker con-figurations: 1) Co-located; target/maskers presented from the center loudspeaker at 0°; 2) Symmetrically separated; target at 0° and maskers symmetrically distributed at ±90°. Spatial release from masking (SRM) was calculated as the difference between the SRTs for these two configurations. Performances with the ciPDA were compared to measures made using the patient's own clinical processors.

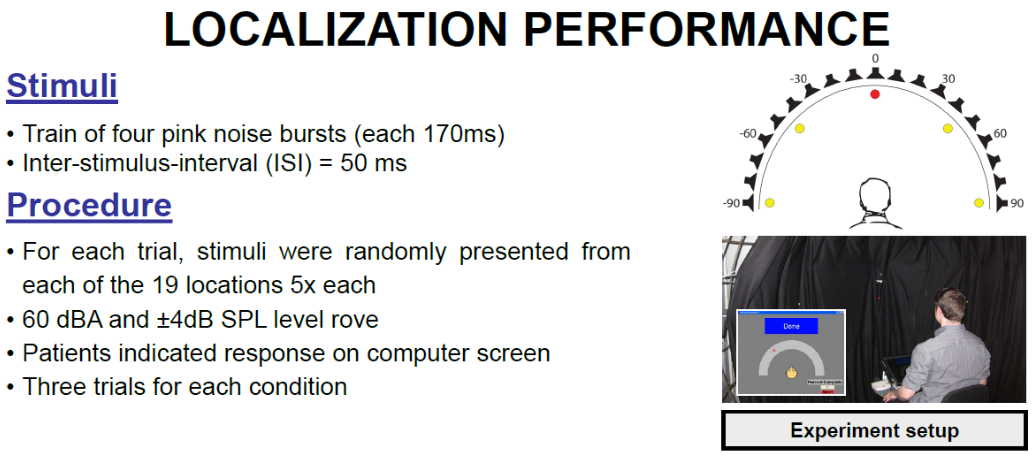


Figure 66: Experimental setup for assessing localization ability with bilateral cochlear implant devices.

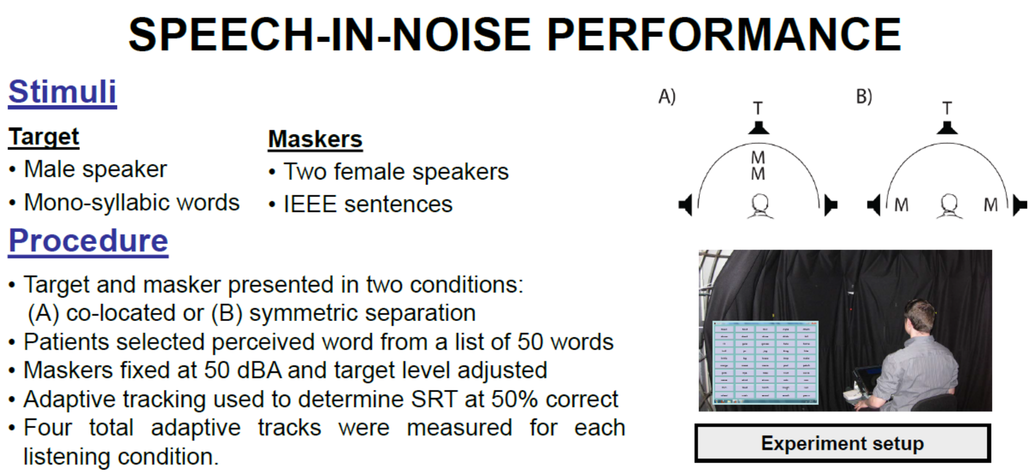


Figure 67: Experimental setup for assessing speech in noise performance with bilateral cochlear implant subjects.

Results showed that acute listening with the ciPDA produced free-field spatial hearing performance comparable to the patient's own clinical processors. The across-subject average RMS errors were nominally lower in the ciPDA listening condition (27.6±3.2°) compared to the clinical condition (32.7±4.2°); however, this difference was not significant. In general, SRTs were lower for the clinical processors compared to the ciPDA. Some listeners exhibited larger SRM benefits (~3-6 dB) using the ciPDA, while others had either larger SRM with their clinical or equivalent SRM between conditions.

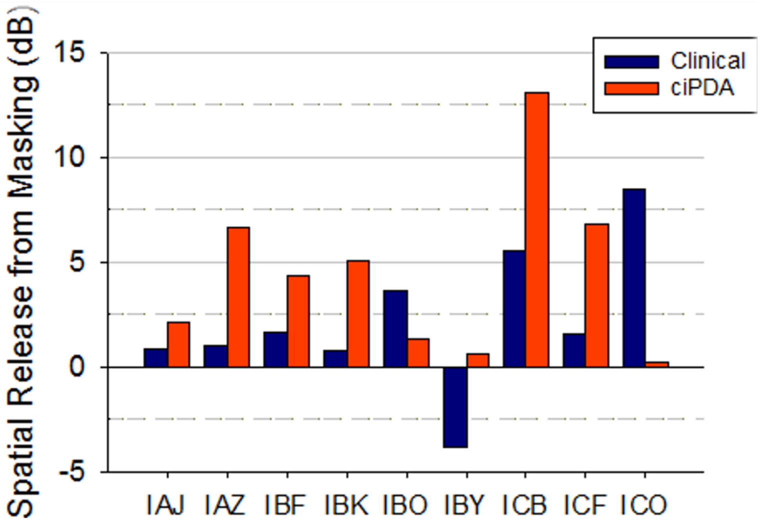


***Figure 68: Compared with conditions in which patients listened with their own clinical processors, acute listening with the ciPDA produced an average of 5-6 dB more SRM**, i.e., greater advantage from spatial separation of the target and competing speech. This finding provides proof of concept for the efficacy of the ciPDA for improving some aspects of binaural processing.*

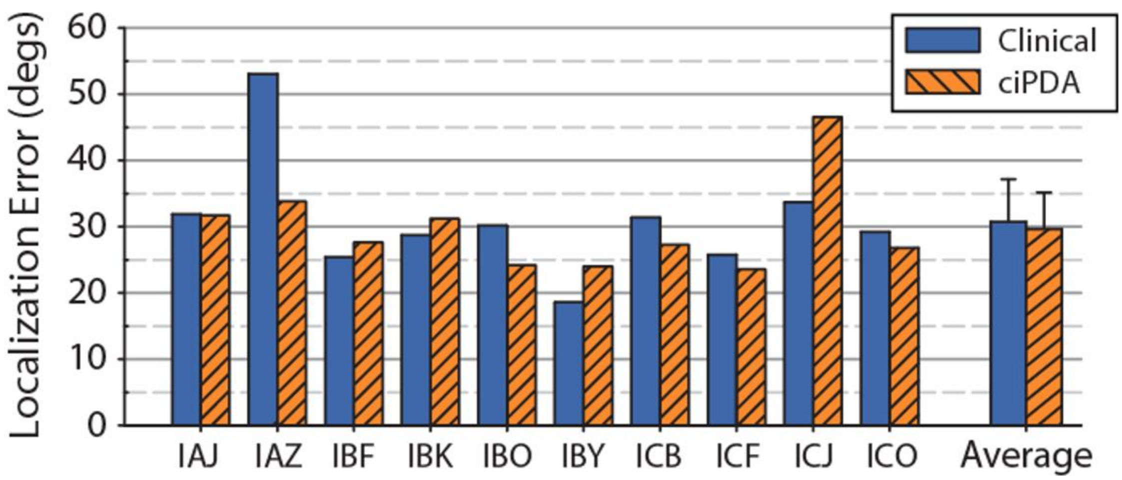
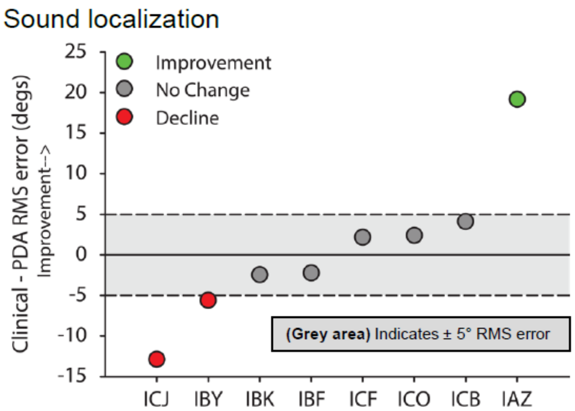


Figure 69: Sound localization performance is shown, and for several subjects errors are lower with the CiPDA than the clinical strategy.



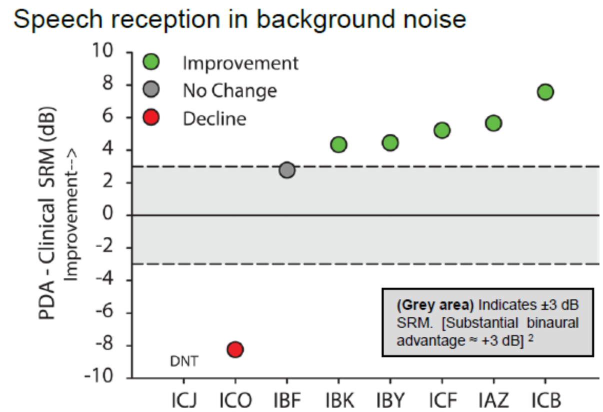


Figure 70: Individual Performance Change - Patients had minimal sound localization improvement, but most exhibited an increase in SRM when listening with the CiPDA

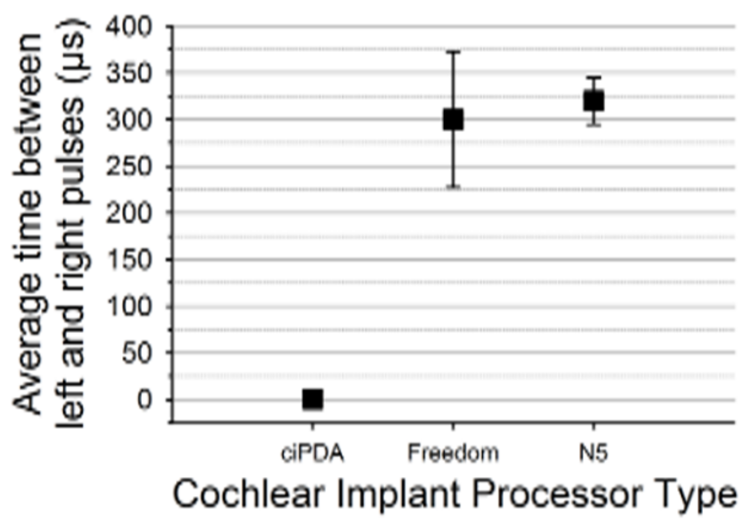


Figure 71: Litovsky lab measured the output of clinical CI processors (N5 or Freedom), and that for CiPDA for stimuli with a 0 ITD, (i.e., right & left processors activated simultaneously). A possible explanation for improvement with ciPDA over conditions with clinical devices is that timing electrical pulses between the two electrode arrays appears to be exactly 0us for the CCi-MOBILE, and to vary on the order of +/-300 us for clinical processors.

Acute listening with a new hardware and software that synchronized stimulation between CIs in the two ears produced benefits in some but not all subjects. However, this work demonstrates that the UTD ciPDA platform provides an effective means for testing novel algorithms and strategies in a CI users. In addition, although currently the device is only used in acute experiments, its success for free-field studies of spatial hearing may provide opportunity for further investigations.

Noise Reduction and Speech Enhancement Schemes for Binaural hearing:
A number of algorithms were proposed that utilized stereo signal principles, e.g., from dual-microphones residing in the same or contralateral ears (hearing-aids or cochlear implants) to reduce noise and improve speech signal quality. These algorithms were based on coherence function, beam forming, and SNR differences in the acoustic signals. These algorithms were evaluated using speech models as well as with groups of normal hearing and hearing impaired individuals. The results indicated great potential for improvements in speech intelligibility with dual-microphones in unilateral and bilateral schemes.

A Dual-Microphone Speech Enhancement Algorithm Based on the Coherence Function: A novel dual-microphone speech enhancement technique was developed. This technique utilized the coherence between the target and noise signals as a criterion for noise reduction and can be generally applied to arrays with closely spaced microphones, where noise captured by the sensors is highly correlated. The proposed algorithm is simple to implement and requires no estimation of noise statistics. In addition, it offers the capability of coping with multiple interfering sources that might be located at different azimuths. The proposed algorithm was evaluated with normal hearing listeners using intelligibility listening tests and compared against a well-established beamforming algorithm. Results indicated large gains in speech intelligibility relative to the baseline (front microphone) algorithm in both single and multiple-noise source scenarios. The proposed algorithm was found to yield substantially higher intelligibility than that obtained by the beamforming algorithm, particularly when multiple noise sources or competing talker(s) were present. Objective quality evaluation of the proposed algorithm also indicated significant quality improvement over that obtained by the beamforming algorithm. The intelligibility and quality benefits observed with the proposed coherence-based algorithm make it an ideal candidate for both hearing aid and cochlear implant devices.

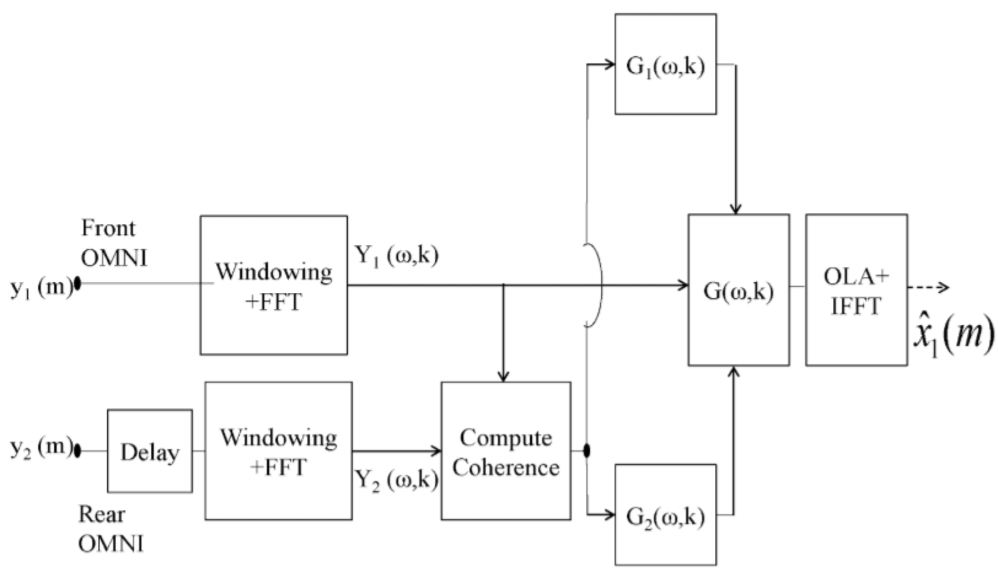


Figure 72: Block diagram of the proposed two-microphone speech enhancement technique.

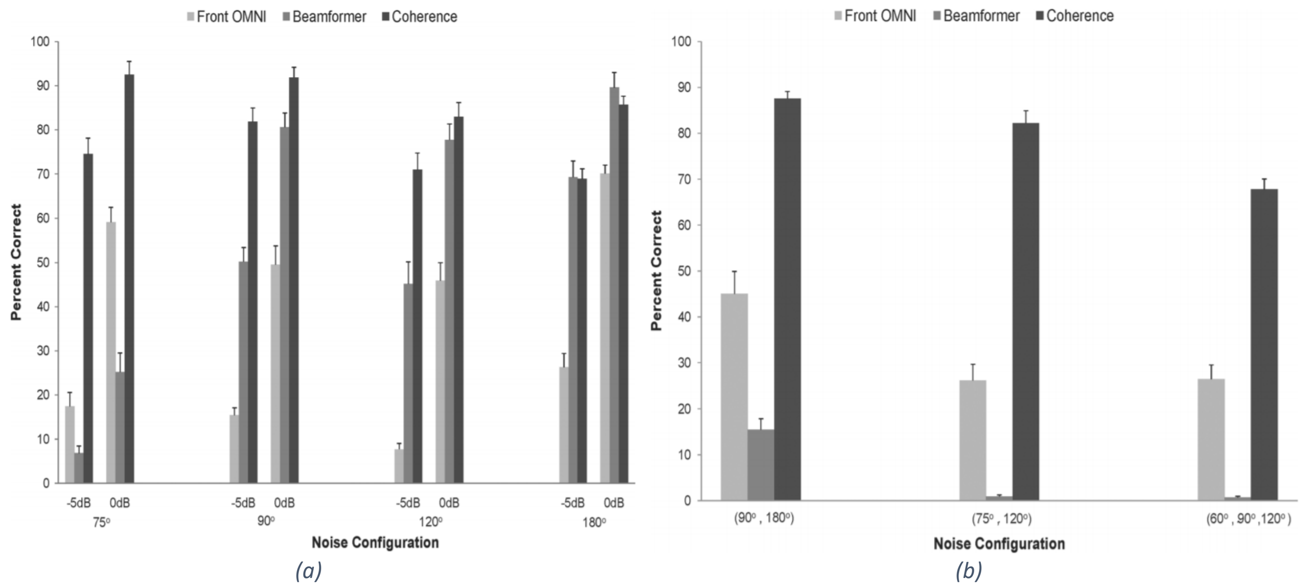


***Figure 73:** (a) Mean percent word recognition scores for ten normal-hearing listeners tested on IEEE sentences in single-noise source scenarios. (b) Mean percent word recognition scores for ten normal-hearing listeners tested on IEEE sentences in multiple-noise sources scenarios (SNR=0dB). Error bars indicate standard deviations*

Further details on the algorithm, methods, and evaluations can be found in: N. Yousefian and P. C. Loizou, "A Dual-Microphone Speech Enhancement Algorithm Based on the Coherence Function," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 20, no. 2, pp. 599-609, Feb. 2012.

A Dual-Microphone Algorithm That Can Cope With Competing-Talker Scenarios: In this work a novel technique was developed for signal-to-noise ratio (SNR) estimation for scenarios where two closely-spaced microphones are available. The proposed technique utilized the real and imaginary parts of the coherence function between the input signals to estimate the SNR without assuming prior knowledge of the noise statistics. The corresponding dual-microphone speech enhancement algorithm utilized a Wiener filter as a gain function constructed using the SNR values computed by the coherence function. Since the proposed SNR estimation technique does not require access to noise statistics, it can be applied in situations where interfering speakers are present. An adaptive speech reception threshold (SRT) test was used to assess the intelligibility of speech processed by the proposed algorithm in scenarios where one or two interfering talkers were present in anechoic and reverberant conditions. Intelligibility listening tests were conducted with both normal-hearing (NH) and cochlear implant (CI) listeners. Results revealed significant improvements in intelligibility and quality over a (baseline) fixed directional algorithm and a well-established beamformer algorithm. In a nearly anechoic room with competing talkers, the improvement in SRT obtained relative to the directional microphone ranged from 5–10 dB, while the improvement obtained by the beamformer was about 2 dB. In reverberant environments, the improvement in SRT remained high (4–7 dB) at T_{60}=220ms, and decreased to 1–2 dB at T_{60}=465ms. Overall, the proposed algorithm provided significant benefits in intelligibility in anechoic and mildly reverberant environments making it suitable for hearing aid and cochlear implant applications.

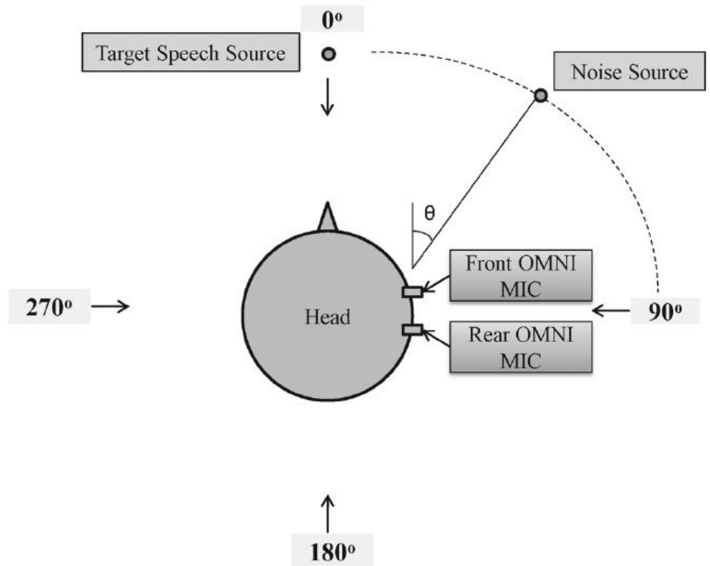


Figure 74: Placement of the two omnidirectional microphones and sound sources.

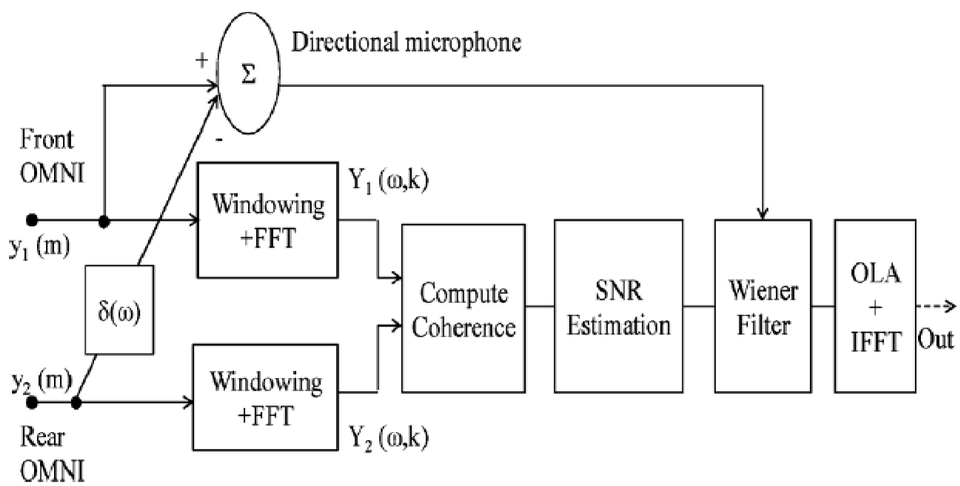


Figure 75: Block diagram of the proposed two-microphone speech enhancement technique.

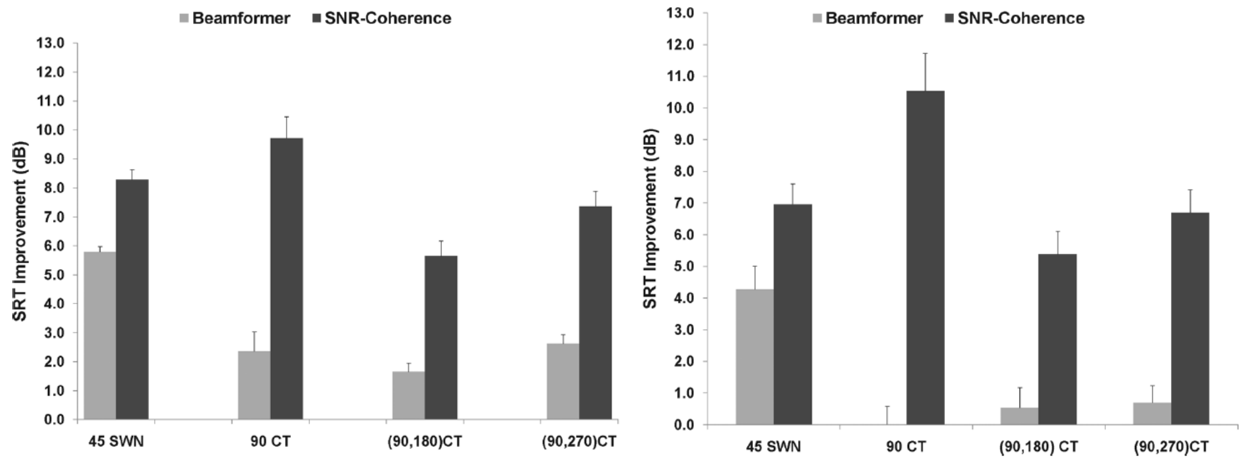


Figure 76: SRT improvements of the beamformer and proposed algorithm over the DIR in different noise configurations, obtained by NH (left panel) and CI (right panel) listeners. Error bars indicate standard deviations.

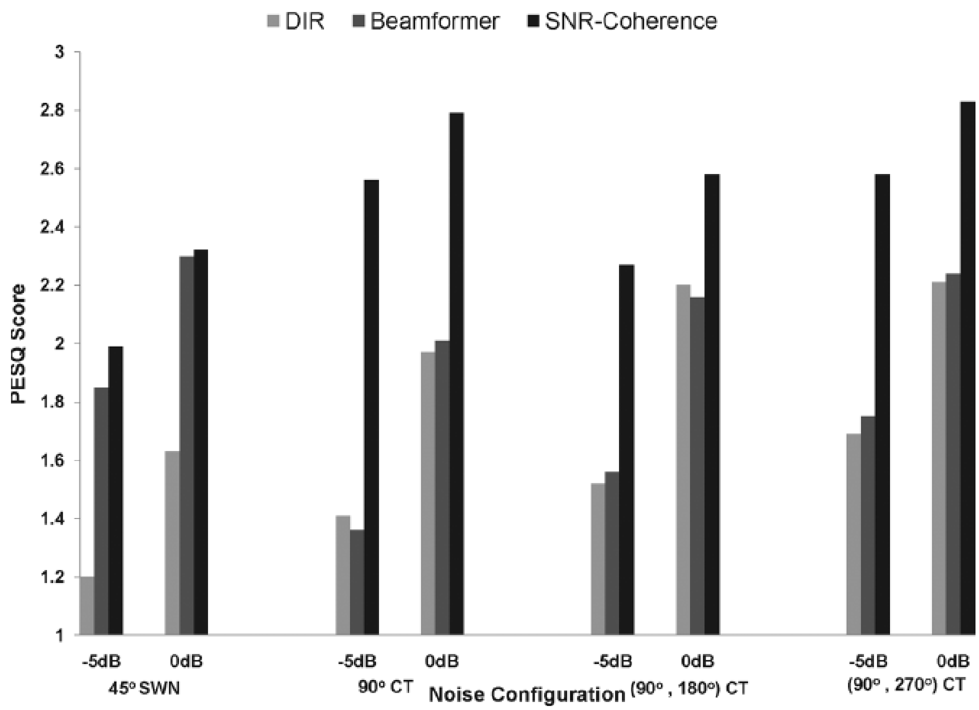


Figure 77: PESQ scores obtained in different noise scenarios.

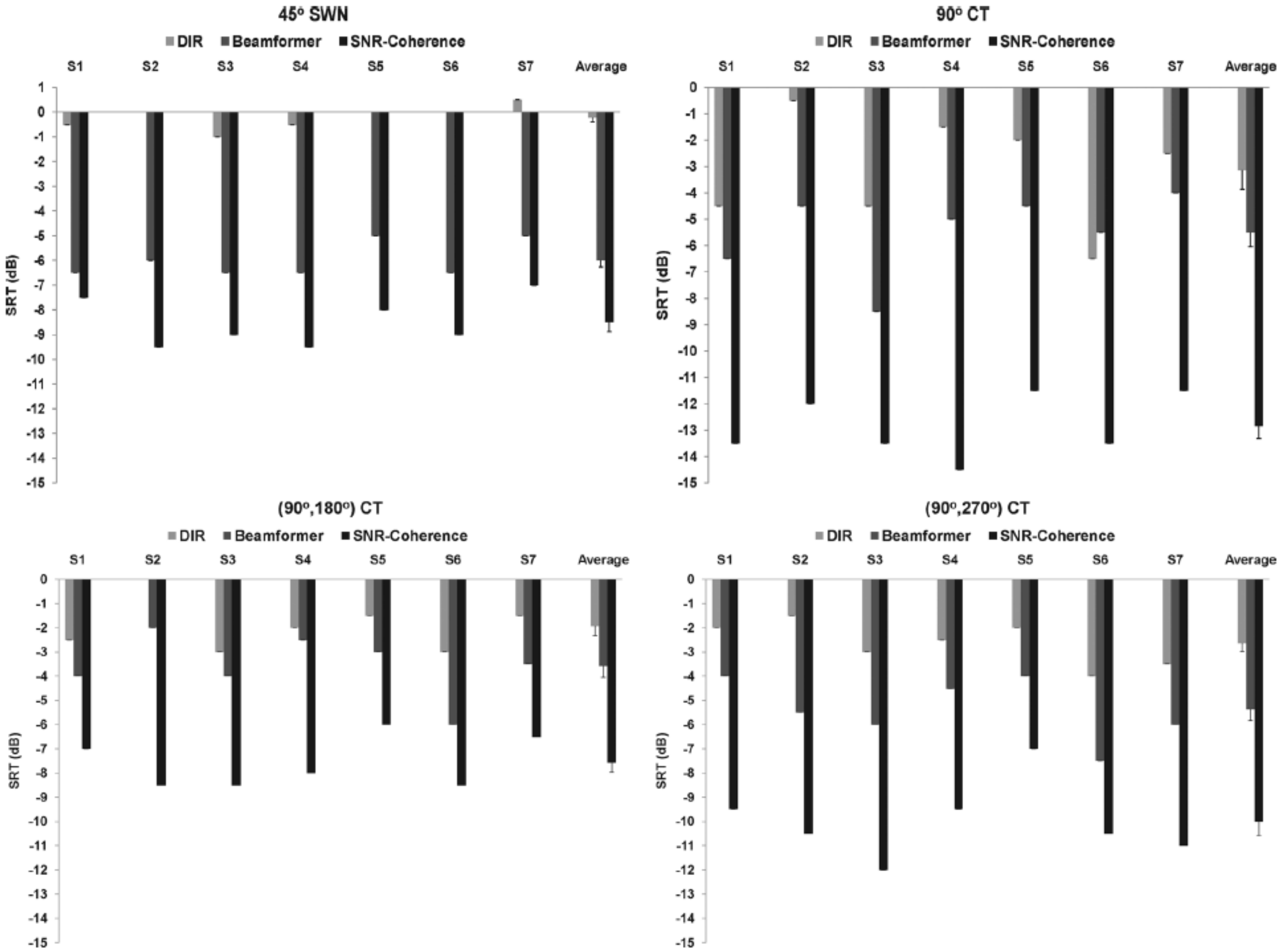
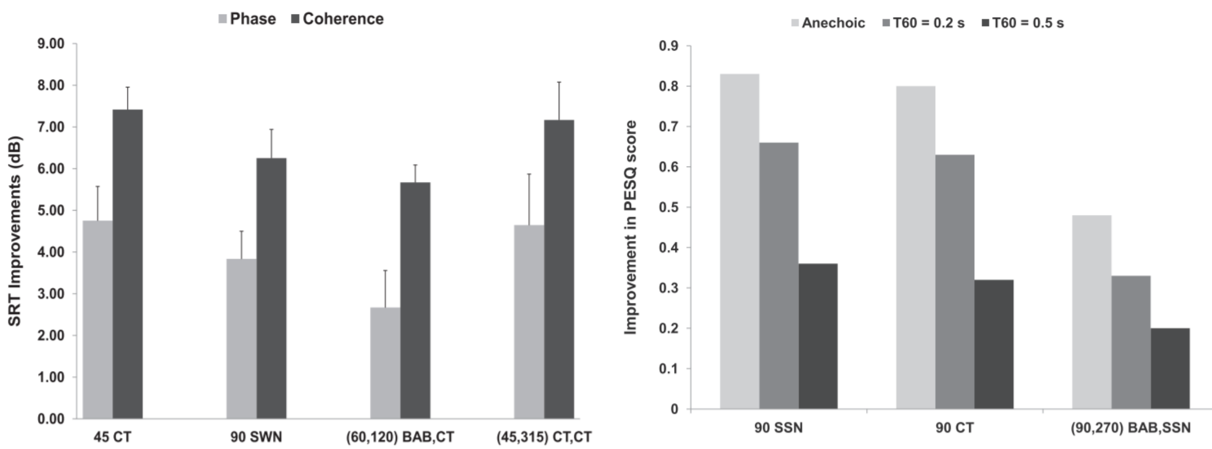


Figure 78: Mean SRT values (dB) obtained by normal-hearing listeners in different noise configurations. Numbers indicate the SNR (dB) needed to understand 50% of the words in sentences correct. Error bars indicate standard deviations.

Further details on the algorithm, methods, and evaluations can be found in: N. Yousefian and P. C. Loizou, "A Dual-Microphone Algorithm That Can Cope With Competing-Talker Scenarios," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 1, pp. 145-155, Jan. 2013.

A coherence-based noise reduction algorithm for binaural hearing aids and cochlear implants: In this worked, a novel coherence-based noise reduction technique was developed and it's use in binaural hearing aids and cochlear implant devices to suppress any potential noise present inside a realistic low reverberant environment was demonstrated. The technique was based on particular assumptions on the spatial properties of the target and undesired interfering signals and suppresses (coherent) interferences without prior statistical knowledge of the noise environment. The proposed algorithm is simple, easy to implement and has the advantage of high performance in coping with adverse signal conditions such as scenarios in which competing talkers are present. The technique was assessed by measurements with normal-hearing subjects and the processed outputs in each ear showed significant improvements in terms of speech intelligibility (measured by an adaptive speech reception threshold (SRT) sentence test) over the unprocessed signals (baseline). In a mildly reverberant room with T 60 = 200, the average improvement in SRT obtained relative to the baseline was approximately 6.5 dB. In addition, the proposed algorithm was found to yield higher intelligibility and quality than those obtained by a well-established interaural time difference (ITD)-based speech enhancement algorithm. These attractive features make the proposed method a potential candidate for future use in commercial hearing aid and cochlear implant devices.



(a) (b)

Figure 79: (a) SRT improvements of the phase based and proposed algorithm over the noisy unprocessed in different noise configurations. Error bars indicate standard deviations, (b) PESQ improvements of the proposed algorithm over the baseline (unprocessed) in three different environments.

Further details on the algorithm, methods, and evaluations can be found in: N. Yousefian, P.C. Loizou, J.H.L. Hansen, "A Coherence-based noise reduction algorithm for binaural hearing aids," Speech Communication, vol. 58, pp. 101-110, Nov. 2013

[Bimodal Cochlear Implants]

Bimodal cochlear implant patients, i.e., patients who have an implant in one ear and low-frequency hearing in the contralateral ear, generally achieve higher scores on tests of speech understanFding than patients who receive stimulation from a cochlear implant only. Experiments were conducted to assess the bimodal benefit in CI/HA population and to determine for bimodal CI users how speech understanding varies as a function of the difference in level between the CI signal and the acoustic signal.

Five postlingually deafened, bimodal CI listeners were invited to participate in this project based on evidence of bimodal benefit in a standard clinical test environment. All provided signed informed consent forms as per institutional guidelines at Arizona State University. All subjects (1) used a Cochlear Corporation signal processor, (2) had at least 24 months of experience with electric stimulation, (3) had at least 20 years of experience with amplification prior to implantation, and (4) were known, from previous testing, to have bimodal benefit.

Signal processing was implemented on CiPDA in an 'offline' mode, i.e. speech signals were processed offline based on each patient's mapping parameters, stimulation parameters for electric stimulation were computed using the ACE algorithm, and the resulting parameters were passed directly to the internal receiver. Direct stimulation removed the need for a sound booth or a quiet room and loudspeakers to deliver speech signals to the patient. The PDA also functioned as a hearing aid and synchronously provided (1) parameters for electrical stimulation and (2) an acoustic signal to the contralateral ear, delivered via an insert receiver, processed in the manner of a hearing aid. To accommodate the degree of hearing loss for each patient, acoustic signals were amplified and subjected to the frequency-gain characteristic prescribed by the National Acoustic Laboratories (NAL-RP) prescription method using MATLAB. The speech stimuli were sentences from the AzBio sentence test lists. The noise signal was multi-talker babble.

The overall logic of the experiment was as follows: (1) present the electric signal in sufficient noise to drive performance near 60% correct and then (2) add (to the electric signal) the acoustic signal at levels ranging from just above detection threshold to above the level of the CI signal. The results for each patient in Figure 80.

The results suggested that the benefit of adding the low-frequency signal to the CI signal is an inverted U-shaped function of the level of the acoustic signal. Very soft acoustic signals (as expected) produce very little benefit: the first data point in each figure is near the level of performance with the CI alone. Signals presented near the balance point between acoustic and electric stimulation always provided the highest levels of performance. And, for all subjects, performance fell for acoustic signals that were judged to be much higher in level than the CI stimulus.

Inspection of Figure 80 suggests that for 3 of the 5 subjects (S1, S4 and S5) an acoustic stimulus that is judged slightly softer, or slightly unbalanced towards the CI ear, provides the most benefit when added to the CI signal. Thus, a clinical recommendation for setting the acoustic signal in balance with, or slightly softer than, the CI signal is supported by the data. Moreover, the data suggest that balance does not need to be determined with a high degree of precision.

Acoustic signals set to MCL, without reference to the CI signal, also support high levels of benefit. When the acoustic signals were set to MCL, scores, on average, were within 4 percentage points of the best score. Thus, balancing is not necessary to achieve a high level of benefit from the acoustic signal. This could be a useful datum for setting the level of the auditory signal for patients who have difficulty understanding CI versus acoustic signal balancing procedures.

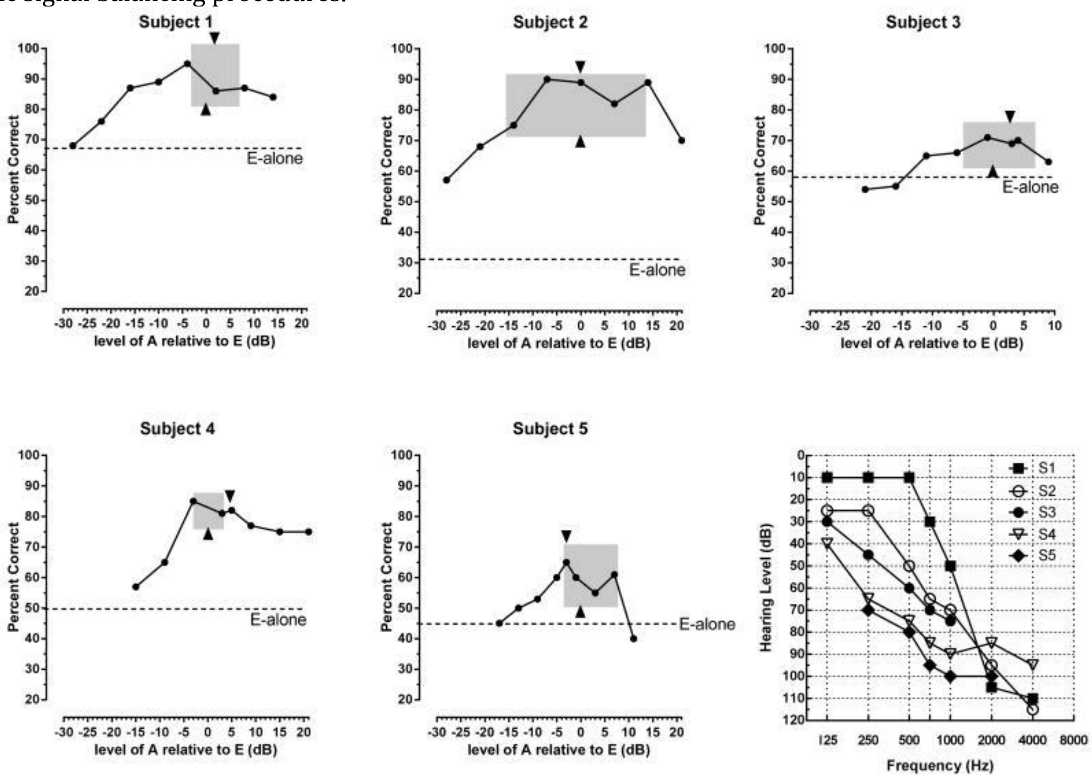


Figure 80: Percent correct sentence understanding as a function of the level of the acoustic signal. E = CI stimulation. A = low-frequency acoustic stimulation. Left edge of grey box = acoustic stimulation level just softer than the CI level. Right edge of grey box = acoustic stimulation level just louder than the CI level. Arrow at bottom of box = balance point for acoustic and electric stimulation. Arrow at top of box = MCL for acoustic stimulation. Bottom right figure shows audiograms of the acoustically stimulated ear.

For all subjects, some signals labeled as softer than the CI signal provided benefit to intelligibility. This suggests that patients who cannot achieve normal loudness growth may still benefit from acoustic stimulation.

The data suggest that (1) acoustic signals perceived as significantly softer than a CI signal can contribute to speech understanding in the bimodal condition, (2) acoustic signals that are slightly softer than, or balanced with, a CI signal provide the largest benefit to speech understanding, and (3) acoustic signals presented at MCL provide nearly as much benefit as signals that have been balanced with a CI signal. Studies involving a larger number of patients and other kinds of test materials are necessary to assess the generality of these results.

[Outreach – STEM education, High School, Undergraduate student engagement]

As part of STEM education outreach and student engagement, an undergraduate bioengineering student team worked with the group for a year to develop middle ear models and signal processing strategies that can prevent harmful effects of very loud impulsive noises. The success of the project lead to showcasing of their work as an educational tool at several meetings/conferences (including Acoustical Society meeting and IEEE Signal Processing Educational Workshop) as well as at the Perot Museum of Science and Technology (Dallas, TX) for yearly Engineer's Week (K-12 STEM outreach). Some details of this effort are provided below.

Educational Electro-Mechanical Model of the Middle Ear:

The overall purpose of this research effort was to design an educational tool and develop a model highlighting the functions of the middle ear bones and the natural safety mechanism of the ear. Illustrating the functionality of the middle ear bones is important, since the middle ear provides some natural level of safety/suppression for excessive levels of sound below 2kHz, but no safety mechanism above 2kHz, which is particularly troubling for either annoying, loud, or extreme high energy impulsive sounds. Also, individuals who have cochlear implants do not have a natural safety mechanism. The scope of this project included development of a standalone, interactive, and educational electro-mechanical model that exhibits the motion of the middle ear bones which include: (i) anatomical 3-bone configuration, (ii) fluid environment in the cochlea, and (iii) electrode stimulation to the auditory nerve cortex. This model has been assessed and approved by STEM/SEEC-(Science and Engineering Education Center) UTDallas.

The two most important parts of the model are the display and information as well as the natural safety mechanism representation. Behind the mechanical model is a display board to help children and kids in a museum-like setting understand the importance behind sound conduction and the need for cochlear implants shown in figure below. Information about sound conduction, the middle ear bones, the eardrum, cochlear implants, the stapes, the cochlea, and the natural safety mechanism of the ear are all explained throughout a K-12 comprehension level. This model can be used simultaneously along instruction or can be used individually by the students themselves. Since the natural safety mechanism of the ears are so important, two springs, one attached to the malleus and one to the incus represent the stapedius and tensor tympanic muscles that contract in order to reduce the mechanical vibrations to protect the movement from damaging the ear.

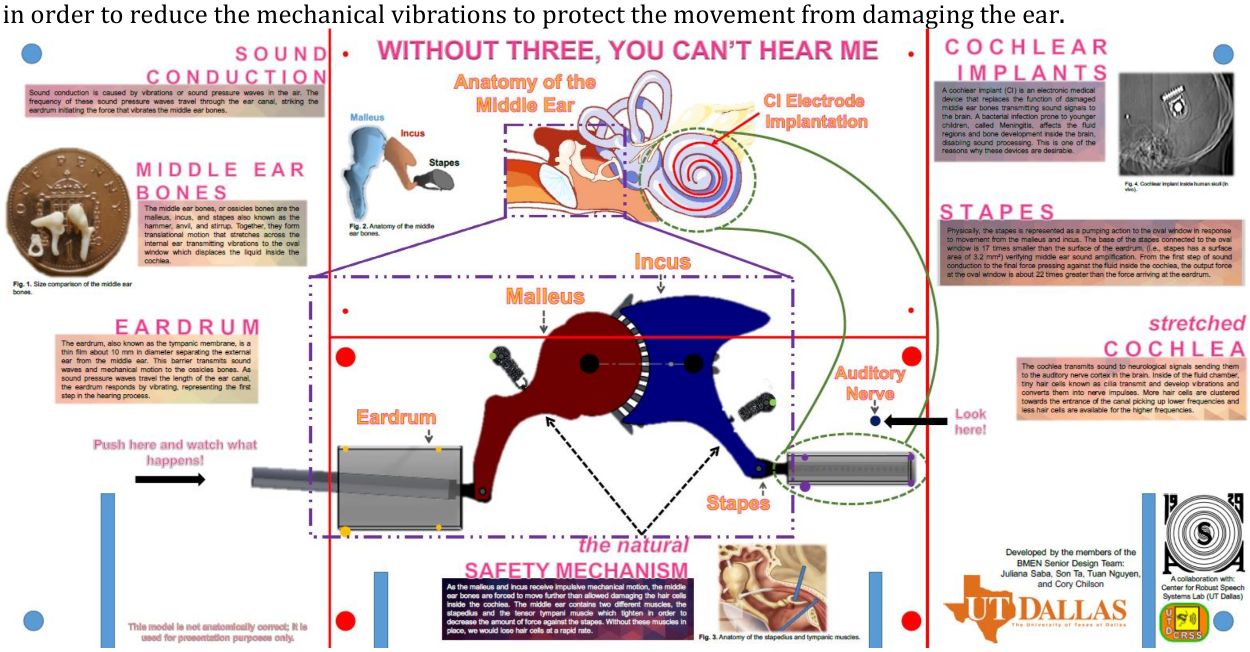


Figure 81: Poster of the middle ear electro-acoustic model ***(presented at Acoustical Society of America, Perot Museum of Science & Technology, and IEEE Signal Processing Education Workshop)***

Figure 82: Undergraduate Bioengineering student team with the 3D model of middle ear. ***(BME Senior Design Team during UTDesign Day Defense (left) and in the CRSS-CILab during a visit from Dr. Blake Wilson (Duke Univ.).***

Perot Museum event for Engineers Week:

CRSS-UTDallas participated in the Ross Perot Museum of Science & Technology (Dallas, TX) during Engineers Week (Feb. 2015, 2016; 2017). The Engineer's daily event had 1300 K-8 grade students attending, and CRSS-UTDallas participated with an interactive 3D model of middle ear. In the United States, approximately 15% of adults (37.5M) age 18 and over report some trouble hearing, 1 in 8 people (13%, or 30M) 12 years or older have hearing loss in both ears, and approximately 3 out of 1000 children are born with hearing loss. The motivation for this effort was to educate the public, especially K-12 students, on the dangers of hearing loss. Having an interactive physical model of the middle ear along with signal processing simulation of effective impulsive sound suppression for hearing aids/cochlear implants helped to provide a useful, hands-on experience for student education. Students had the opportunity to interact with the model, along with talking with UTDallas UG students and CRSS faculty/staff.

Acoustical Society Meeting:

The educational electro-mechanical model of the middle ear was showcased in Fall 2014 Acoustical Society of America (ASA) meeting. The model was appreciated a lot and the team was asked to build two additional models as part of ASA's outreach efforts for educating about hearing mechanisms. Further details about the model and work can be found in the publication below:

Juliana Saba, Hussnain Ali, Jaewook Lee, John H.L. Hansen, Son Ta, Tuan Nguyen, and Cory Chilson, "Development of an educational electro-mechanical model of the middle ear," Proceedings of the 170th (Fall) meeting of the Acoustical Society of America, JASA vol. 138 (3), Nov. 3, 2015.

IEEE Signal Processing and SP Education Workshop:

IEEE Signal Processing and SP Education Workshop (SP/SPE) was held in Salt Lake City, UT in 2015. The educational electro-mechanical model of the middle ear built by the bioengineering UG team was also showcased at the workshop. In addition to the model, impulsive-type noise suppression algorithm and its performance with actual CI user was also reported. Further details on this can be found in:

Juliana Saba, Son Ta, Tuan Nguyen, Cory Chilson, Jaewook Lee, Hussnain Ali, and John H.L. Hansen, "Developing an educational electro-mechanical model of the middle ear and impulse noise reduction algorithm for cochlear implant users," IEEE Signal Processing and SP Education Workshop (SP/SPE), Salt Lake City, UT, 2015, pp. 83-88.

Figure 83: CRSS-UTDallas participated in the Ross Perot Museum of Science & Technology (Dallas, TX) during Engineers Week (Feb. 2015, 2016; 2017); their demonstration of electro-mechanical middle-ear model as well as signal processing capabilities for impulse and noise suppression for hearing aids/cochlear implants were highlighted.

Key Outcomes or Other Achievements:

[Mobile research platform for cochlear implant and hearing aid research]

- Have successfully designed and tested a PDA-based research platform for CI and HA research.
- Successfully migrated the platform to newer generation of Android smartphones/tablets.
- Platform can successfully work in both realtime and offline modes with a PC and an Android phone/tablet.
- Developed software tools, libraries and applications (in MATLAB, C, and JAVA) that enable conducting custom experiments with the platform.
- Successfully tested the platform for reliable output and safe stimulation by conducting electrical reliability tests as well as by conducting several human experiments. No adverse effects were experienced in any of the experiments.

[Reverberation and noise suppression algorithms for cochlear implants]

- Analyzed the role of combinations of noise and reverberation on speech recognition performance with cochlear implant devices.
- Developed at least six different signal processing algorithms that could mitigate noise, reverberation, and combination of both.
- Evaluated the performance of algorithms with groups of CI recipients and found significant improvements in speech recognition and speech quality with the proposed/developed schemes.

[Speech enhancement strategies for improved speech intelligibility and quality]

- Developed noise-robust F0 estimation strategies that ultimately lead to the development of novel speech enhancement methods for cochlear implant users.
- Evaluated the performance of speech enhancement strategies with CI users. The data from experiments indicated significant gains in speech intelligibility with the developed techniques.
- The developed techniques for pitch estimation yielded substantially better performance than the compared state-of-the-art algorithms across various types and levels of noise.

[Physiologically inspired algorithms for improved speech recognition with CI devices in everyday realistic environments]

- Systematically studied the effect of environmental noise on speech production of cochlear implant users and for the first time provided scientific evidence on the presence of Lombard Effect in the speech produced by post-lingually deafened cochlear implant users.
- Developed a Lombard corpus by collecting speech data from CI users. Systematically studied and presented perceptual analysis of Lombard effect in speech of CI users.
- Developed a novel Lombard perturbation strategy that was inspired by speech production/feedback of CI users. Experiments with CI study participants indicated significant gains in speech understanding and quality with the developed algorithms.

[Customized sound processing and fitting strategies for cochlear implants]

- Developed custom image-guided tailor-fit sound processing and fitting strategies for cochlear implant users. This first of its kind patient-centric signal processing strategy was developed in collaboration with Vanderbilt University Medical Center.
- Evaluated the custom sound processing strategy with groups of normal hearing and cochlear implant users acutely, semi-acutely, and chronically using take-home trials.
- The outcomes from the studies indicated quicker adaptation and higher asymptotic level of performance with the proposed strategy.

[Bilateral cochlear implants]

- Systematically studied sound localization performance in bilateral cochlear implant users with synchronized electrical stimulation in both ears.
- Systematically studied speech-in-noise performance in bilateral cochlear implant users with synchronized electrical stimulation in both ears.
- The outcomes from the studies indicated improvements in speech reception in competing background noise with synchronized electrical stimulation.

[Bimodal listening]

- Scientifically assess the bimodal benefit in CI/HA population and to determine for bimodal CI users how speech understanding varies as a function of the difference in level between the CI signal and the acoustic signal.
- Experiments with bimodal CI/HA subjects provided clarifications on the levels of acoustic input and the resulting bimodal benefits in terms of gains in speech recognition performance.

[Other achievements]

- Developed objective speech intelligibility measures for cochlear implant users in complex listening environments.
- Conducted chimeric experiments to investigate the perceptual importance of temporal envelop and temporal fine structure between tonal and non-tonal language.
- Quantized the effect of consonant-vowel contributions to speech perception in cochlear implants.
- Studied amplitude modulation detection with cochlear implants.

What opportunities for training and professional development has the project provided?

- Team members have worked collaboratively with research teams from several universities and research centers including Arizona State University, University of Wisconsin-Madison, New York University, University of Colorodo Boulder, University of Southern California, Vanderbilt University and others.
- Cochlear implant research community had a special interest in the research platform developed by our laboratory in this project. This allowed us to establish collaborative research projects with external university and industry partners.

How have the results been disseminated to communities of interest?

- A number of primary and secondary publications have resulted from the research advancements (summarized later in this report). The publications that describe the project, research problems and solutions have been published, and are available to researchers worldwide.
- Workshops and trainings were conducted throughout the project lifecycle at conferences and meetings to disseminate latest R&D updates, availability, and features of the research platform. Some of these events were:
  - First UT-Dallas PDA Workshop in December 2010 at UT-Dallas.
  - Update Meeting during 36th MidWinter ARO meeting on February 2013 at Baltimore, MD.
  - UT-Dallas Resrach platform updates during CIAP 2013 in July 2013 at Lake Tahoe, CA.
  - U-Dallas platform meeting during CIAP 2015 in July, 2015 at Lake Tahoe, CA.
  - UT-Dallas research platform workshop during CIAP 2017 in July 2017 at Lake Tahoe CA.

- The Univ. of Texas – Dallas Principal Investigators (Loizou, Tobey, and Hansen) has given a number of seminars based on project advancements, as well as participated in outreach opportunities for K-12 and the community.
- The research platform and tools have been distributed free of cost to several research laboratories and centers worldwide. Names and affiliations of research groups who have received the platform (as a result of this project) are given below:

	Investigator	Platform version
1.	**Dr. Michael Dorman** Cochlear Implant Research Laboratory, Arizona State University, Email: MICHAEL.DORMAN@asu.edu Phone: (480)965-3345	CI PDA platform v 2.1.0 Serial # DOR121410
2.	**Dr. Ruth Litovsky** Binaural Hearing and Speech Lab, The University of Wisconsin at Madison, Email: litovsky@Waisman.Wisc.Edu Phone: (608) 262 5045	CI PDA platform v 2.1.0 Serial # LIT121410 CCiMOBILE-v1 Serial # UWMAW16152CFC

3.	**Dr. Bomjun Kwon** Department of Hearing, Speech and Language, Gallaudet University, Email: bomjun.kwon@gallaudet.edu	CI PDA platform v 2.1.0 Serial # KWO121410
4.	**Dr. Leslie Collins** Duke University Email: lcollins@ee.duke.edu Phone: 919-660-5260	CI PDA platform v 2.1.0 Serial # LES102310
5.	**Dr. Mario Svirsky** NYU Langone Medical Center New York University Email: mario.svirsky@nyumc.org Phone: (212) 263-7217	CI PDA platform v 2.1.0 Serial # SVI121410 CCiMOBILE-v1 Serial # NYUAU1612887B
6.	**Dr. Lina Reiss** Oregon Health & Science University Email: reiss@ohsu.edu	CI PDA platform v 2.1.0 Serial # REI111813
7.	**Dr. Christopher Brown** University of Pittsburgh Email: cbrown1@pitt.edu	CI PDA platform v 1 Serial # Board14
8.	**Dr. Jay Rubenstein and Dr. Kaibao Nie** University of Washington Email: rubinj@u.washington.edu Phone: 206-616-6655	CI PDA platform v 1 Serial # RUB121410
9.	**Dr. Roger Miller** NIDCD, NIH Email: millerr@nidcd.nih.gov	CI PDA platform v 1 Serial # LOTCKX
10.	**Dr. Ray Goldsworthy** House Ear Institute Email: raygold@sens.com	CI PDA platform v 1 Serial # RAY120110
11.	**Lakshmish Ramanna** Cochlear Corp. Denver Email: lramanna@cochlear.com	CI PDA platform v 1 Serial # LAK021011
12.	**Dr. Bas Van Dijk** Global Research Coordinator - Sound Coding Cochlear Technology Centre Belgium Email: BVanDijk@cochlear.com	CCIMobile-v1 Serial # BCOAU1612887A
13.	**Dr. Julio Cordioli** Laboratório de Vibrações e Acústica Universidade Federal de Santa Catarina, Brazil Email: julio.cordioli@ufsc.br	CCIMobile-v1 Serial # FUSAU16122CFD
14.	**Dr. Fan-Gang Zeng** Hearing and Speech Lab University of California – Irvine Email: fzeng@uci.edu	CI PDA platform v 1 Serial # ZEN121410

Products

You have the option of selecting "nothing to report" in this section. There are no limitations to the number of entries you submit and you can also pull information directly from Thomson Search when using the online tool on Research.gov.

Within the Products section, you can list any products resulting from your project during the specified reporting period, such as:

Technologies:

1. PDA-based (Ci-PDA) research platform for cochlear implant and hearing-aid research
2. Android-based (CCi-MOBILE) research platform for cochlear implant and hearing aid research
3. Software, sound processing scripts and codes, Apps for sound processing and custom experimental needs with cochlear implants.

Thesis/Dissertations:

Doctorate Theses:

[Thesis1] Oldooz Hazrati, "Development of dereverberation algorithms for improved speech intelligibility by cochlear implant users,"Ph.D. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, 2012.

[Thesis2] Nima Yousefian "Dual-Microphone and binaural Noise Reduction Techniques for Improved Speech Intelligibility for Hearing Aid Users," Ph.D. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, May 2013.

[Thesis3] Hussnain Ali, "Customized sound processing and fitting paradigms for cochlear implant users," Ph.D. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, May 2016.

[Thesis4] Dongmei Wang, "Speech Analysis and Single Channel Enhancement to Improve Speech Intelligibility for Cochlear Implant Recipients," Ph.D. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, March 2017.

[Thesis5] Jaewook Lee, "Lombard Effect in Speech Production by Cochlear Implant Users: Analysis, Assessment, and Implications," Ph.D. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, March 2017.

[Thesis6] Hua Xing, "Advancements in Statistical Signal Processing and Machine Learning for Speech Enhancement," Ph.D. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, March 2017.

M.S. Theses:

[Thesis7] Rithika Saripella, "Impact of selective amplification of consonants on speech intelligibility in noise by hearing impaired listeners," M.S. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, Dec. 2010.

[Thesis8] Nageswara Rao Gunupudi, "Adaptive psychoacoustic experiments for cochlear implants using the PDA research platform," M.S. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, Dec. 2010.

[Thesis9] Milad Omidi, "The contribution of syllabic information to speech intelligibility in quiet and in noise," M.S. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, May 2012.

[Thesis10] Hussnain Ali, "Design and evaluation of a PDA-based research platform for cochlear implant," M.S. Thesis, Cochlear Implant Laboratory, The University of Texas at Dallas, May 2012.

Tutorials:

[Tutorial1] IEEE ICASSP-2015 (April 2015): Tutorial Speaker: "Signal Processing for Cochlear Implants," (Brett Swanson, Oldooz Hazrati, Michael Goorevich, John H.L. Hansen) IEEE ICASSP-2015, Brisbane, Australia, April 19-24, 2015.

[Tutorial2] ISCA Interspeech-2016 (Sept. 2016): Tutorial Speaker: "Hearing Assistive Technologies: Challenges and Opportunities," (Oldooz Hazrati, Hussnain Ali, Jim Kates, John H.L. Hansen), Interspeech-2016, San Francisco, CA (USA), Sept. 8, 2016.

Journals:

Primary supported peer-reviewed publications:

[JP1] D.Wang, J.H.L. Hansen, "Single Channel Speech Enhancement based on Harmonic Estimation combined with Statistical based Approach to Improve Speech Intelligibility for Cochlear Implant Recipient," Journal of the Acoustical Society of America, submitted Feb. 2017. Revised Sept. 2017.

[JP2] D. Wang, C. Yu, J.H.L.Hansen, "Robust Harmonic Features for Classification based Pitch Estimation," IEEE Trans. Audio, Speech, and Language Processing, vol. 25, no. 5, pp. 952-964, May. 2017.

[JP3] J. Lee, H. Ali, A. Ziaei, E. Tobey, J.H.L. Hansen, "The Lombard Effect observed in Speech produced by Cochlear Implant users in Noisy Environments: A Naturalistic Study," Journal of the Acoustical Society of America, vol. 141, no. 4, pp. 2788-2799, April, 2017.

[JP4] H. Xing, J.H.L. Hansen, "Single Sideband Frequency Offset Estimation and Correction for Quality Enhancement and Speaker Recognition," IEEE Trans. Audio, Speech, and Language Processing, vol. 25, no. 1, pp. 124-136, Jan. 2017.

[JP5] O. Hazrati, H. Ali, J.H.L. Hansen, E. Tobey, "Evaluation and analysis of whispered speech for cochlear implant users: Gender identification and intelligibility," Journal of the Acoustical Society of America, vol. 138, No. 1, 74-79, July 2015.

[JP6] N. Yousefian, J.H.L. Hansen, P. Loizou, "A Hybrid Coherence Model for Noise Reduction in Reverberant Environments," IEEE Signal Processing Letters, vol. 22, no. 3, pp. 274-277, Mar. 2015.

[JP7] H. Ali, O. Hazrati, E. Tobey, J.H.L. Hansen, "Evaluation of adaptive dynamic range optimization in adverse listening conditions for cochlear implants," Journal of the Acoustical Society of America - JASA Express Letters, vol. 136, no. 3, pp. 1515-1528, Sept. 2014.

[JP8] C. Yu, K.K. Wójcicki, P.C. Loizou, J.H.L. Hansen, "Evaluation of the importance of time-frequency contributions to speech intelligibility in noise," Journal of the Acoustical Society of America, vol. 135, (5), pp. 3007-3016, 2014.

[JP9] T.H. Falk, V. Parsa, J.F. Santos, K. Arehart, O. Hazrati, R. Huber, J. Kates, and S. Scollie, "Objective Quality and Intelligibility Prediction for Users of Assistive Listening Devices", IEEE Signal Processing Magazine, Special Issue for Signal Processing Techniques for Assisted Listening, Vol. 32, No. 2, pp. 114-124, March 2015

[JP10] N. Yousefian, P.C. Loizou, J.H.L. Hansen, "A Coherence-based noise reduction algorithm for binaural hearing aids," Speech Communication, vol. 58, pp. 101-110, Nov. 2013

[JP22] H. Ali, A. Lobo, P. Loizou, "Design and Evaluation of a Personal Digital Assistant-based Research Platform for Cochlear Implants," IEEE Transactions on Biomedical Engineering (TBME), vol. 60, issue 11, pp. 3060 – 3073, Nov. 2013

[JP12] O. Hazrati, O. Sadjadi, P.C. Loizou, J.H.L. Hansen, "Simultaneous suppression of noise and reverberation in cochlear implants using a ratio masking strategy," Journal of the Acoustical Society of America, Vol. 134, No. 5, pp. 3759-3765, Nov. 2013.

[JP13] J.F. Santos, S. Cosentino, O. Hazrati, P.C. Loizou, and T.H. Falk, "Objective Speech Intelligibility Measurement for Cochlear Implant Users in Complex Listening Environments", Speech Communication, 55(7), 815-824, 2013.

[JP14] O. Hazrati, and P.C. Loizou, "Reverberation Suppression in Cochlear Implants Using a Blind Channel-Selection Strategy", J. Acoust. Soc. Am., 133(6), 4188-4196, 2013.

[JP15] O. Hazrati, and P.C. Loizou, "Comparison of Two Channel Selection Criteria for Noise Suppression in Cochlear Implants", J. Acoust. Soc. Am., 133(3), 1615-1624, 2013.

[JP16] O. Hazrati, J. Lee, and P.C. Loizou, "Blind Binary Masking for Reverberation Suppression in Cochlear Implants", J. Acoust. Soc. Am., 133(3), 1607-1614, 2013.

[JP17] N. Yousefian and P. C. Loizou, "A Dual-Microphone Algorithm That Can Cope With Competing-Talker Scenarios," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 1, pp. 145-155, Jan. 2013.

[JP18] F. Chen, O. Hazrati, and P.C. Loizou, "Predicting the Intelligibility of Reverberant Speech for Cochlear Implant Listeners with. Non-intrusive Intelligibility Measure", Biomedical Signal Processing & Control, 8(3), 311-314, 2013.

[JP19] N. Yousefian and P. C. Loizou, "A Dual-Microphone Speech Enhancement Algorithm Based on the Coherence Function," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 20, no. 2, pp. 599-609, Feb. 2012.

[JP20] N. Yousefian, P.C. Loizou, "Predicting the speech reception threshold of cochlear implant listeners using an envelope-correlation based measure," The Journal of the Acoustical Society of America 2012 132:5, 3399-3405.

[JP21] KK Wójcicki, PC Loizou, "Channel selection in the modulation domain for improved speech intelligibility in noise," The Journal of the Acoustical Society of America, vol. 131 (4), pp. 2904-2913, 2012.

[JP22] O. Hazrati, and P.C. Loizou, "The Combined Effect of Reverberation and Noise on Speech Intelligibility by Cochlear Implant Listeners", Int. J. Audiol., 51(6), 437-443, 2012.

[JP23] O. Hazrati, and P.C. Loizou, "Tackling the Combined Effects of Reverberation and Masking Noise using Ideal Channel Selection", J. Speech. Lang. Hear. Res., 55, 500-510, 2012.

[JP24] Kostas Kokkinakis, Oldooz Hazrati, and Philipos C. Loizou, "A Channel-Selection Criterion for Suppressing Reverberation in Cochlear Implants", J. Acoust. Soc. Am., 129(5), 3221-3232, 2011.

Secondary supported peer-reviewed publications, which contribute to the project:

Books: [nothing to report]

Book Chapters: [nothing to report]

Conference Papers and Presentations:

Primary supported peer-reviewed publications:

[CP1] D. Wang, J.H.L. Hansen, "Speech Enhancement Based on Harmonic Estimation combined with MMSE to Improve Speech Intelligibility for Cochlear Implant Recipients," ISCA INTERSPEECH-2017, paper#78, pp. 186-190, Stockholm, Sweden, Aug. 20-24, 2017.

[CP2] Hussnain Ali, Ammula Sandeep, Juliana Saba, John H.L. Hansen, "CCi-MOBILE platform for cochlear implant and hearing-aid research," Proc. Of the 1st Int. Conference on Challenges in Hearing Assistive Technology (CHAT-17), Stockholm, Sweden, Aug. 19, 2017.

[CP3] Rafael Chiea, Bernardo Murta, Gustavo Mourao, Stephan Paul, Hussnain Ali, John Hansen, "Towards a measure of the differences in cochlear implant stimulation strategies," Proc. Of the 1st Int. Conference on Challenges in Hearing Assistive Technology (CHAT-17), Stockholm, Sweden, Aug. 19, 2017.

[CP4] Hussnain Ali and John H.L. Hansen, "Internet-of-Things and smart assistive hearing devices," Conf. on Implantable Auditory Prostheses, CIAP-2017, pp. 212, Lake Tahoe, CA, July 16-21, 2017.

[CP5] J. Lee, H. Ali, J.H.L. Hansen, "Lombard Effect perturbation pre-processing strategy for cochlear implant users," CIAP-2017: Conf. on Implantable Auditory Prostheses, Lake Tahoe, CA, July 16-21, 2017.

[CP6] Hussnain Ali, Sandeep Ammula, John H.L. Hansen, "Subjective evaluation with UT-Dallas research interface for cochlear implant users," Conf. on Implantable Auditory Prostheses, CIAP-2017, pp. 211, Lake Tahoe, CA, July 16-21, 2017.

[CP7] Julian N. Saba, Hussnain Ali, John H.L. Hansen, "Improving channel selection criteria in n-of-m strategies for cochlear implant sound coding strategies," Conf. on Implantable Auditory Prostheses, CIAP-2017, pp. 227, Lake Tahoe, CA, July 16-21, 2017.

[CP8] Sandeep Ammula, Hussnain Ali, John H.L. Hansen, "CCi-MOBILE platform for combined electric and acoustic stimulation," Conf. on Implantable Auditory Prostheses, CIAP-2017, pp. 261, Lake Tahoe, CA, July 16-21, 2017.

[CP9] D. Wang, J.H.L. Hansen, "Comparing speech intelligibility benefits of F0-informed speech enhancement with cochlear implant listeners between stationary noise and non-stationary noise," CIAP-2017, Conf. on Implantable Auditory Prostheses, Lake Tahoe, CA, July 16-21, 2017.

[CP10] D. Wang, J.H.L. Hansen," Pitch pattern matching based speech enhancement ," The Acoustical Society of America, 4pSPb, Topics in Signal Processing in Acoustics, Boston, MA, June 25-29, 2017.

[CP11] D. Wang, J.H.L. Hansen,"Single channel speech enhancement based on harmonic estimation combined with MMSE framework to improve speech intelligibility for cochlear implant recipient," The Acoustical Society of America, 5aSPa, Audio and Array Signal Processing, Boston, MA, June 25-29,2017.

[CP12] J. Lee, H. Ali, J.H.L. Hansen, "Intelligibility Enhancement of Neutral Speech based on Lombard Effect Modification with Application to Cochlear Implant Users," ARO-2017: Association for Research Otolaryngology, Baltimore, MD, Feb. 11-15, 2017.

[CP13] H. Xing, G. Liu, J.H.L. Hansen, "Frequency Offset Correction in Single Sideband (SSB) Speech based on Deep Neural Networks for Speaker Verification," ISCA INTERSPEECH-2015, Paper#503, pp. 1156-1160,

Dresden, Germany, Sept. 6-10, 2015.

[CP14] J. Saba, S. Ta, T. Nguyen, C. Chilson, J.W. Lee, H. Ali, J.H.L. Hansen, "Development of an Educational Electro-Mechanical Model of the Middle Ear," Acoustical Society of America, Salt Lake City, USA, 23-27, 2016.

[CP15] D. Wang, J.H.L. Hansen, "Single channel speech enhancement to improve the speech intelligibility for cochlear implant user," 14th Inter. Conf. Cochlear Implants and other Implantable Auditory Technologies, CI-2016, paper 2016-A-782-ACI (poster talk), Toronto, Canada, May 11-14, 2016.

[CP16] Hussnain Ali, Feng Hong, Jun Wang, John H.L. Hansen, "Mobile research interface for cochlear implants," 14th International Conference on Cochlear Implants and Other Implantable Technologies, CI2016, Toronto, Canada, May 11 - May 14, 2016.

[CP17] J.W. Lee, H. Ali, J.H.L. Hansen, "Effect of Lombard Speech on Speech Intelligibility in Adult Cochlear Implant Users," 14th Inter. Conf. Cochlear Implants and other Implantable Auditory Technologies 2016, paper 2016-A-890-ACI (poster talk), Toronto, Canada, May 11-14, 2016.

[CP18] Hussnain Ali, Jack H. Noble, René H. Gifford, Robert F. Labadie, Benoit M. Dawant, John H.L. Hansen, "Towards patient-centric sound fitting for cochlear implants," 14th International Conference on Cochlear Implants and Other Implantable Technologies, CI2016, Toronto, Canada, May 11 - May 14, 2016.

[CP19] Juliana Saba, Jaewook Lee, Hussnain Ali, Son Ta, Tuan Nguyen, and John Hansen, "Impulse suppression algorithm development of a compatible program for cochlear implant users," Proceedings of the 170th (Fall) meeting of the Acoustical Society of America, JASA vol. 139 (4), pp. 2224, April 27, 2016.

[CP20] D. Wang, J.H.L. Hansen, "F0 Estimation for Noisy Speech by exploring Temporal Harmonic Structures in Local Time-Frequency Spectrum Segments" IEEE ICASSP-2016: Inter. Conf. on Acoustics, Speech, and Signal Processing, paper #3752, pp. 6510-6514, Shanghai, China, March 20-25, 2016.

[CP21] D. Wang, J.H.L. Hansen, "F0 Estimation for Noisy Speech based on Exploring Local Time-Frequency Segments", IEEE WASPAA-2015. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, Mohonk Mountain House, New Paltz, NY, Oct. 18-21, 2015

[CP22] J.N. Saba, H. Ali, J.W. Lee, J.H.L. Hansen, S. Ta, T. Nguyen, C. Chilson, "Developing an Educational Electro-Mechanical Model of the Middle Ear," 170th Meeting of the Acoustical Society of America, 138, no. 3, 2aED12, pp. 1772, Sept. 2015 (presentation Nov. 3, 2015).

[CP23] J.N. Saba, S. Ta, T. Nguyen, C. Chilson, J. Lee, H. Ali, J.H. L. Hansen, "Developing an Educational Electro-Mechanical Model of the Middle Ear and Impulse Noise Reduction Algorithm for Cochlear Implant Users," IEEE Signal Processing Society Education Workshop, Paper#M1-13, pp. 431-436, Snowbird, Utah, Aug. 9-12, 2015.

[CP24] Hussnain Ali, Jack H. Noble, René H. Gifford, Robert F. Labadie, Benoit M. Dawant, John H.L. Hansen, Emily Tobey, "Image-Guided Frequency-Place Mapping in Cochlear Implants," Conference on Implantable Auditory Prostheses, CIAP-2015, pp. 183, Lake Tahoe, CA, July 12-17, 2015.

[CP25] Feng Hong, Hussnain Ali, John H.L. Hansen, E. Tobey, "Android-based Research Platform for Cochlear Implants," Conf. on Implantable Auditory Prostheses, CIAP-2015, pp. 127, Lake Tahoe, CA, July 12-17, 2015.

[CP26] D. Wang, J.H.L. Hansen, E. Tobey, "Speech enhancement based on glimpse detection to improve the

speech intelligibility for cochlear implant recipient," CIAP-2015: Conf. on Implantable Auditory Prostheses, Lake Tahoe, CA, July 12-17, 2015.

[CP27] J.W. Lee, H. Ali, A. Ziaei, J.H.L. Hansen, E.A. Tobey, "Impact analysis of naturalistic environmental noise type on speech production for cochlear implant users versus normal hearing listeners," CIAP Conf. on Implantable Auditory Prostheses, Lake Tahoe, CA, July 12-17, 2015.

[CP28] J. Lee, H. Ali, A. Ziaei, J.H.L. Hansen, "Analysis of Speech and Language Communication for Cochlear Implant Users in Noisy Lombard Conditions," IEEE ICASSP-2015, Paper#4016, pp. 5132-5136, Brisbane, Australia, April 19-24, 2015.

[CP29] Oldooz Hazrati, Shabnam Ghaffarzadegan, and John H. L. Hansen, "Leveraging Automatic Speech Recognition in Cochlear Implants for Improved Speech Intelligibility Under Reverberation", Proc. IEEE ICASSP'15, Brisbane, Australia, April 2015.

[CP30] Hussnain Ali, Jack H. Noble, René H. Gifford, Robert F. Labadie, Benoit M. Dawant, John H.L. Hansen, Emily Tobey, "Image-guided customization of frequency-place mapping in cochlear implants," Proceedings of International Conference on Acoustics, Speech, and Signal Processing, IEEE ICASSP'15, pp. 5843 – 5847, Brisbane, Australia, April 2015. (Invited paper)

[CP31] J. Lee, H. Ali, A. Ziaei, J.H.L. Hansen, "Lombard Effect based Speech Analysis across Noisy Environments for Voice Communications with Cochlear Implant Subjects," Acoustical Society of America Indianapolis, IN, Oct. 2014.

[CP32] D. Wang, J. Kates, J.H.L. Hansen, "Investigation of the relative perceptual importance for temporal envelop and temporal fine structure between tonal language and non-tonal language," ISCA Interspeech-2014, Paper #1087, Singapore, Sept. 14-18, 2014

[CP33] D. Wang, P.C. Loizou, J.H.L. Hansen, "F0 estimation for noisy speech based on long-term harmonic feature analysis combined with neural network classification," ISCA Interspeech-2014, Paper #1093, Singapore, Sept. 14-18, 2014

[CP34] D. Wang, P.C. Loizou, J.H.L. Hansen, "Noisy Speech Enhancement based on Long Term Harmonic Model to Improve Speech Intelligibility for Hearing Impaired Listeners," ISCA Interspeech-2014, Paper #1107, Singapore, Sept. 14-18, 2014

[CP35] Hussnain Ali, Oldooz Hazrati, John H. L. Hansen, and Emily Tobey, "The Effect of Adap-tive Dynamic Range Optimization on Speech Intelligibility in Adverse Listening Environments for Cochlear Implant Users", in Proc. 13th International Conference on Cochlear Implants and Other Implantable Auditory Technologies, Munich, Germany, June 2014.

[CP36] Oldooz Hazrati, Hussnain Ali, John H. L. Hansen, and Emily Tobey, "Gender Identification and Intelligibility of Whispered Speech in Cochlear Implant Users: Evaluation and Analysis", in Proc. 13th International Conference on Cochlear Implants and Other Implantable Auditory Technologies, Munich, Germany, June 2014.

[CP37] Hussnain Ali, Oldooz Hazrati, John H. L. Hansen, and Emily Tobey, "The effect of adaptive dynamic range optimization on speech intelligibility in adverse listening environments for cochlear implant users," in Proceedings of 13th International Conference on Cochlear Implants and Other Implantable Auditory Technologies, CI 2014, Munich, Germany, June 2014.

[CP38] Oldooz Hazrati, Hussnain Ali, John H. L. Hansen, and Emily Tobey, "Gender identification and intelligibility of whispered speech in cochlear implant users: evaluation and analysis," in Proc. 13th International Conference on Cochlear Implants and Other Implantable Auditory Technologies, CI 2014, Munich, Germany, June 2014.

[CP39] Hussnain Ali, Feng Hong, John H. L. Hansen, and Emily Tobey, "Improving channel selection of sound coding algorithms in cochlear implants," in Proc. International Conference on Acoustics, Speech, and Signal Processing, IEEE ICASSP'14, pp. 905 - 909, Florence, Italy, May 2014.

[CP40] Oldooz Hazrati, Seyed Omid Sadjadi, and John H. L. Hansen, "Robust and Efficient En-vironment Detection for Adaptive Speech Enhancement in Cochlear Implants", in Proc. IEEE ICASSP'14, Florence, Italy, May 2014

[CP41] H. Xing, J.H.L. Hansen, "Frequency offset correction in single sideband speech for speaker verification," IEEE ICASSP-2014, IEEE Inter. Conf. on Acoustics, Speech and Signal Processing, pp. 4050-4054, (paper #1569858861), Florence, Italy, May 4-9, 2014.

[CP42] Oldooz Hazrati, John H.L. Hansen, and Emily Tobey, "Efficient Environment Detection for Adaptive Speech Enhancement in Cochlear Implants", in Proc. Association for Research in Oto-laryngology'14, San Diego, California, USA, February 2014.

[CP43] Oldooz Hazrati, Seyed Omid Sadjadi, Hussnain Ali, Philipos C. Loizou, and John H. L. Hansen, "A Soft Masking Strategy for Simultaneous Suppression of Noise and Reverberation in Cochlear Implants", in Proc. Conference on Implantable Auditory Prostheses'13, Lake Tahoe, California, USA, July 2013.

[CP44] Nirmal Srinivasan, Hussnain Ali, Emily Tobey, Philipos C. Loizou, "Effect of consonant-vowel boundary to speech perception in cochlear implants," Conference on Implantable Auditory Prostheses '13, (CIAP 2013) Lake Tahoe, California, USA, July 2013.

[CP45] Feng Hong, Hussnain Ali, Emily Tobey, and Philipos C. Loizou, "A flexible monopolar stimulator for animal studies in auditory prostheses," Conference on Implantable Auditory Prostheses '13, (CIAP 2013) Lake Tahoe, California, USA, July 2013.

[CP46] D. Wang, P. Loizou, J.H.L. Hansen, "Monaural Speech Enhancement for Cochlear Implants based on Pitch Estimation combined with CASA Binary Mask Estimation," CIAP-2013: Conf. on Implantable Auditory Prostheses, Paper R36, Lake Tahoe, CA, July 14-19, 2013.

[CP47] C. Yu, K. Wojcicki, P. Loizou, J.H.L. Hansen, "A New Mask-Based Objective Measure for Predicting the Intelligibility of Binary Masked Speech," IEEE ICASSP-2013, IEEE Inter. Conf. on Acoustics, Speech and Signal Processing, pp. 7030-7033, (paper #5002), Vancouver, Canada, May 26-31, 2013

[CP48] Nirmal Srinivasan, Oldooz Hazrati, and Philipos C. Loizou, "The Contribution of Room Expo-sure and Semantic Context to Speech Intelligibility in Cochlear Implant Users", in Proc. Association for Research in Otolaryngology'13, Baltimore, Maryland, USA, February 2013.

[CP49] Hussnain Ali, Oldooz Hazrati, Nirmal Srinivasan, and Philipos C. Loizou, "Noise Modeling in Rerberation: A Comparative Study of Speech Intelligibility in Cochlear Implant Users", in Proc. Sigma Xi, Dallas, Texas, USA, January 2013.

[CP50] O. Hazrati, J Lee and P.C. Loizou, "Binar.y Mask Estimation for Improved Speech Intelligibility in Reverberant Environments", in Proc. INTERSPEECH'12, Portland, Oregon, USA, September 2012.

[CP 51] Joao Felipe Santos, Stefano Cosentino, Oldooz Hazrati, Philipos C. Loizou, and Tiago H. Falk, "Performance Comparison of Intrusive Objective Speech Intelligibility and Quality Metrics for Cochlear Implant Users", in Proc. INTERSPEECH'12, Portland, Oregon, USA, September 2012.

[CP52] Hussnain Ali, Arthur Lobo, Philip Loizou, "On the Design and Evaluation of the PDA-based Research Platform for Electric and Acoustic Stimulation", in Proc. 34th Annual International Conf. of Engineering in Medicine and Biology Society IEEE EMBS, pp. 2493-2496, San Diego, USA, Aug. 2012. (Invited Paper)

[CP 53] Oldooz Hazrati, Jaewook Lee and Philipos C. Loizou, "The Contribution of Vowel/Consonant Boundaries to Speech Recognition in Reverberation by Cochlear Implant Users", in Proc. Association for Research in Otolaryngology'12, San Diego, California, USA, February 2012.

[CP54] Hussnain Ali, Arthur Lobo, Philip Loizou, "PDA Platform for Offline Processing and Streaming of Stimuli for Cochlear Implants Research," 33rd Annual International Conf. of Engineering in Medicine and Biology Society IEEE EMBS, pp. 1045-1048, Boston, USA, Sep. 2011.

[CP 55] Oldooz Hazrati, Kostas Kokkinakis, and Philipos C. Loizou, "Suppressing Reverberation in Cochlear Implants Using a Channel-Selection Based Strategy," in Proc. Conference on Implantable Auditory Prostheses'11, Pacific Grove, California, USA, July 2011.

[CP 56] Philipos C. Loizou, Oldooz Hazrati, Nima Yousefian, Kostas Kokkinakis, and Fei Chen, "Channel Selection: A Panacea for the Background Interference Problem in Cochlear Implants", in Proc. Conference on Implantable Audiotory Prostheses'11, Pacific Grove, California, USA, July 2011.

[CP 57] Kostas Kokkinakis, Oldooz Hazrati, and Philipos C. Loizou, "A New Channel-Selection Criterion for Better Speech Recognition in Reverberation by Cochlear Implant Users," in Proc. Association for Research in Otolaryngology'11, Baltimore, Maryland, USA, February 2011.

[CP58] N. Yousefian, K. Kokkinakis, PC Loizou, "A coherence-based algorithm for noise reduction in dual-microphone applications," 18th European Signal Processing Conference, 2010.

[CP 59] Oldooz Hazrati, Kostas Kokkinakis, and Philipos C. Loizou, "A Blind Subband-based Dereverberation Algorithm", in Proc. IEEE ICASSP'10, Dallas, TX, USA, March 2010.

Secondary supported peer-reviewed publications, which contribute to the project:

Other Publications:

Patents: [nothing to report]

Inventions: [nothing to report]

Oldooz Hazrati, and Philipos C. Loizou, "Binary Mask Estimation for Enhancement of Reverberant Speech", UTD TechID 11-029.

Licenses: [nothing to report]

Websites: [nothing to report]

Other Products: [nothing to report]

NOTE: You may upload PDF files with images, tables, charts, or other graphics in support of the Products section. You may upload up to 4 PDF files with a maximum file size of 5 MB each.

Participants

There are no limits on the number of participants you list for this section; however, you must list participants who have worked one person month or more for the project reporting period. You have the option of selecting "nothing to report" in this section. The online service will also ask for additional information on participants such as:

- What individuals have worked on the project?
- What organizations have been involved as partners?
- Have other collaborators or contacts been involved?

What individuals have worked on the project?

Name	Most Senior Project Role
Philip C. Loizou	PI
Emily Tobey	PI
John H.L. Hansen	PI and Technical Lead since July 2012.
Ruth Litovsky	Collaborator
Michael Dorman	Collaborator
Hussnain Ali	PhD student and Staff:
Feng Hong	Post-doctorate researcher
Kamil Wojcicki	Post-doctorate researcher
Oldooz Hazrati	PhD Student and Post-doctorate researcher
Nima Yousefian	PhD Student
Nirmal Srinivasan	Post-doctorate researcher
Dongmei Wang	PhD Student
Jaewook Lee	PhD Student
Hua Xing	PhD Student
Juliana Saba	PhD Student
Ammula Sandeep	PhD Student
Milad Omidi	M.S. Student
Vanishree Gopalakrishna	PhD Student
Taher Shahbazi Mirzahasanloo	PhD Student

What other organizations have been involved as partners?

Subcontracts:

1. Ruth Litovsky from University of Wisconsin – Madison
2. Michael Dorman from Arizona State University

Have other collaborators or contacts been involved? **Yes** No

Impacts

You have the option of selecting “nothing to report” in this section.

What is the impact on the development of the principal discipline(s) of the project?

Cochlear prosthesis is widely accepted as the most effective clinical intervention to restore auditory function of individuals with profound hearing loss. The cochlear implant system has been continuously improved from front-ending sound processing and fitting software to internal stimulator and electrode design. In particular, sound processor technology has played a significant role in the growth of CI uptake in the community. The research work conducted during this project focused on the development of new generation of speech processing strategies for unilateral, bilateral and bimodal cochlear implant users that could allow user-customization. In order to achieve these goals a general purpose sound processing research platform was developed which enabled us to design new experiments and evaluate user performance over time. Implant manufacturers usually provide research speech processors for use with human subjects that allow researchers to develop and test new signal processing algorithms. However, most labs are unable to use them due to limited technical resources or due to the constrained framework of the interface provided by the manufacturer. These limitations include flexibility, portability, wearability, ease of programmability, long-term evaluation and features to design intricate experiments. One of the important factors which hinder their use for speech processing research is that a skilled programmer is required to implement the algorithms in a high level or low level language. These limitations/needs were addressed by the research platform developed during this project, which provided a flexible software driven solution for both clinicians and researchers without requiring advanced programming skills or major hardware investment. The platform for the first time provided many unique features in one single package, e.g., synchronized bilateral stimulation, electric+acoustic stimulation, portability, programming flexibility, and numerous software applications. These features have enabled the platform to emerge as a successful tool in the research arena with many researchers utilizing it to implement new and existing algorithms and establishing its use in clinical studies.

The earlier generation of the platform (PDA-based) was the first of its kind research interface which was built on a commercially available PDA device. It set an example of the bridging of medical devices with consumer electronics. Later version of the platform provided more computing resources, software flexibility, and ease of use. The current Android-based version is one-of-a-kind research platform, which is orders of magnitude more flexible and computationally powerful than existing clinical processors and will aid in bridging scientific research with commercial applications. During the course of this project, the platform was shared with the CI research community free of cost using an open source model. The motivation behind open-source and free distribution was to provide research community with the needed technology and research tools to realize their scientific ideas and test their hypothesis and eventually advance the field. The ability to perform real-life chronic speech assessments with the platform is likely to open new frontiers for scientific exploration in future and will result in a paradigm shift in how speech processing/perception research is carried out in the cochlear implant field.

The speech processing advancements brought forth by this project have (and will further) advance the field. Most of the peer-reviewed work from the group has been widely cited by other research groups, which is a testament to the quality and impact of research conducted during this project. Speech enhancement algorithms for noise reduction, de-reverberation, and user-customization developed during this project were all novel formulations and extensive human assessments were conducted to test the performance of these research ideas. There is a huge potential in these schemes to improve speech recognition and speech quality with cochlear implant devices in everyday realistic noisy environments. These schemes could potentially be migrated to clinical cochlear implant systems and bring practical benefit to implant users in their daily lives.

What is the impact on other disciplines?

Nothing to report.

What is the impact on the development of human resources?

The project has supported ten Ph.D. electrical engineering graduate students, as well as two M.S. electrical engineering students. In addition, five post-doctorate researchers were supported by the grant on different aspects of this project. At least three Senior Design Teams (a total of 12 Undergraduate Students) have been participating in Senior Design on the project. Teams have participated in the Biennial Senior Design Day competition, with extensive members from industry involved in formal oral and poster evaluations. Lab alumni were eventually recruited by top tier academic institutes and industry.

What is the impact on physical resources that form infrastructure?

Nothing to report.

What is the impact on institutional resources that form infrastructure?

Nothing to report.

What is the impact on society beyond science and technology?

More than 45 million Americans have a significant hearing loss, and 50% of adults who are 75 years or older have disabling hearing loss. Census projections indicate the number of Americans older than 65 years will increase 157% by the year 2050. Hearing loss is the most common birth defect in developed countries, and 3.5 per 1000 children have permanent significant hearing loss, which can lead to significant impairment in language development and ultimately in educational outcomes. Although the number of implant recipients has grown exponentially over the last two decades, speech recognition performance, has only seen modest gains relative to the pace of actual digital signal processing (DSP) technology. Innovations brought forth by this project in the form of a portable research platform for CI research and novel speech processing strategies that aim at improving speech perception in adverse listening conditions (i.e., in the presence of noise, reverberation, competing talkers, etc.) will have direct benefit to CI listeners. The platform provides unique opportunities to support sophisticated signal processing and stimulation strategies in real-time in everyday environments. It can help test the effectiveness of research ideas and scientific hypotheses in real-life settings, and can potentially assist in advancing current laboratory research to commercial applications. Scientific studies brought forth by this project are likely to provide key insights in the hearing mechanisms of cochlear implant users in everyday listening environments. Also, the algorithms formulated during this project will help CI users in achieving better speech recognition with their devices and ultimately have a better quality of life.

Changes / Problems

If not previously reported in writing to the agency through other mechanisms, provide the following additional information or state, "Nothing to Report", if applicable.

Changes in approach and reason for change:

Nothing to report.

Actual or Anticipated problems or delays and actions or plans to resolve them:

There have been challenges with the robustness and wearability of the research platform for conducting long-term experiments in the field. In order to ensure the platform provided equivalent speech reception performance to the clinical counterpart, extensive evaluations were conducted. These evaluations consisted of benchtop electrical testing as well as human speech assessments. For non-human testing, output stimuli in response to multiple acoustic inputs was captured over long periods of time (in a stress-test paradigm) to ensure correct outputs were received and real-time operation was maintained over the long run. Amplitude and frequency calibrations at different acoustic levels were conducted for each unit. Electrodograms generated by complex acoustic signals (e.g., speech) were compared against clinical implementations and validated for correctness. Finally, cochlear implant subjects were recruited to participate in listening studies to assess speech recognition performance with the platform as well as speech quality. In order to improve the wearability of the device, 3D enclosures were manufactured that could pack the complete platform as one whole unit and subjects could carry the platform in a bag in everyday environments.

Changes that have a significant impact on expenditures:

Nothing to report.

Significant changes in use or care of human subjects:

Nothing to report.

Significant changes in use or care of vertebrate animals:

Nothing to report.

Significant changes in use or care of biohazards:

Nothing to report.

Special Requirements

This report section is only available when Special Requirements are specifically noted in the solicitation and approved by the Office of Management and Budget.

NOTE: You may upload PDF files in support of the Special Requirements section. You may upload PDF files with a maximum file size of 10 MB each. There is no limit to the number of files uploaded.