SYDE 252 Project Phase 1:
Cochlear Implant Signal Processing

Fall 2024, University of Waterloo

Gurnirmal Kaur, 21027178
Niki Shemirani, 21026573
Valerie Stoica, 21016929
Sarah Qiao, 20997346

Introduction

In our evaluation of various cochlear implant designs, we seek to not only emphasize the engineering aspect, but also the individualized psychological aspect of the finalized design to ensure that the sound output is functional, reliable, and pleasant to the user. Because this implant requires effectively irreversible surgical changes, it is important for the design and alternatives to be properly evaluated and reviewed for user preference. For example, many child patients that receive cochlear implants will depend on them to develop their speech and will need to use the devices for up to 30 years before replacing the implant, if at all [1]. These factors impose serious design requirements because flaws can greatly affect the person's daily living and ability to communicate freely.

In the standardized practice of evaluating implant sound output, acoustic simulation is used to allow a normal-hearing listener to understand the sound of the electric simulation and identify key performance parameters. The difficulty in this method is the variability among cochlear implant users; moreover, the qualitative characteristics of the implant are difficult to simulate [1]. Thus, it is important to note this limitation of acoustical simulation and factor in subjective user data, such as through surveys, in our evaluation method. Following recovery from surgery, patients will typically follow up with an audiologist to adjust device settings (e.g. loudness scaling) that are more individualized ear qualities that change based on factors such as age at implantation and length of deafness [2]. Thus, it is also important to factor in the ability to adjust the final system design.

Another notable trait that we are hoping to explore and improve is music perception for patients with cochlear implants. It is known that cochlear implants are not good at conveying the complexities of pitches seen in music. Because of the limitation of 20 electrodes designed in the cochlear implant, it is very difficult to recreate the sounds of a fully functioning auditory system [3]. As the designers of this project, we would like to explore the improvement of this gap and include additional parameters that cover this topic such as frequency range.

Overall, our evaluation method will utilize the standard practice of acoustic simulation, so that we can compare the processed sound with our normal perception of sound. Notably, we will collect various perspectives from third-parties with normal hearing to factor in their subjective hearing preferences (e.g. nuances like clarity and loudness [4]). This is an aspect that can often be overlooked because it is difficult to measure, especially as the scale of distribution for cochlear implants for a company increases, and user preferences become more varied. To factor in these differences, we plan to distribute a survey to assess the more subjective qualities of sound. It is also important to note that, because we are creating a design proposal, we do not have the resources to conduct a real trial or survey with prospective and current cochlear implant patients, which should be heavily considered in real design. Finally, we will rank the processor designs based on the average feedback in a decision matrix to choose an overall top candidate.

Metrics and Evaluation Methodology

The performance of the signal processor will be evaluated based on five key metrics that reflect the user's experience. These metrics include clarity, word intelligibility/similarity to input, adaptability, volume range, and frequency range. These metrics are essential for assessing how well the processor accommodates various real-world scenarios.

1. Clarity
Clarity refers to the *distinctness* of the audio signal output. It is an indication of how free the output is from distortion or noise. High clarity is essential for user comfort, especially in environments where multiple audio sources compete for attention.

Subjective Listening Test:
Participants will listen to various types of sound, processed by the different designs for signal processors. The output audio will be compared to the original input sound. The clarity of each output audio will be rated on a scale of 1 (unclear) to 5 (extremely clear). The average of all ratings will be the score for each design.
Sound Files: Files used will consist of speech samples (male (1.0), female (1.1), multiple speakers(1.2)), instrumental piano (2.0), and environmental sounds (background noise from environments such as classrooms (3.0), and streets (3.1)).

2. Word Intelligibility/Similarity to Input
Word intelligibility is a measure of how well a listener can comprehend and distinguish between words in the signal processor output. In sound samples where there are no spoken words, the similarity of the output sound to the input will be measured. That is, how well the processed signal resembles the original input in terms of tone and sound quality.

Subjective Listening Test:
Participants will listen to a variety of sounds files (multiple speakers, male/female alarms, etc.) and rate their ability to understand words, or their perception of how similar the processed sound is to the original input. If the sound file contains speech, participants will be asked to write down what they hear from the audio sample. This will be compared to the correct transcription of the input file. The design with the highest sum of correct words from all participant trials will be given a score of 5. All other designs will be scored proportionately. If the audio sample does not contain speech, participants will rate the similarity of the processed audio to the input audio on a scale of 1 (very different) to 5 (very similar). The average of all ratings will be used as the final value.
Sounds Files: Files used will consist of various speech samples (male (1.0), female (1.1)), environmental sounds (classrooms (3.0), streets (3.1)). Other sound samples used in this section will include impulse samples such as an alarm clock (4.0) and squash ball sound (4.1).

3. Adaptability
Adaptability refers to the processor's ability to deliver *consistent* performance across a range of environmental conditions, such as varying noise levels, or multiple competing speakers. This is crucial for users who move between different environments.

Subjective Listening Test:
Participants will listen to a variety of sounds that contain changing conditions/environments. The processor's performance will be evaluated in terms of its ability to maintain the same level of clarity and intelligibility across variations in the sound sample; how the processor handles *changes* in

environment. Participants will rate adaptability on a scale of 1 (poor adaptability) to 5 (excellent adaptability). The average of all ratings will be used as the final value.
Sound Files: Files will focus on changing conditions and environments. For example, sound files that consist of talking loudly to whispering (5.0), one speaker to multiple speakers (5.1).

4. Volume Range
Volume range refers to the range of audio levels from soft to loud. It is a measure of how well the processor can handle all audio levels without losing detail or introducing distortion. A broad volume range is important for ensuring users have a clear listening experience, regardless of the volume of the input sound.

Subjective Listening Test:
Audio samples that consist of various volume levels will be played, and participants will rate how well the processor manages these different volumes. Participants will focus on whether both quiet and loud sounds are clearly audible. Based on the performance of the processor on multiple sound files, the volume range will be rated on a scale of 1 (significant distortion or loss of detail at low or high volumes) to 5 (clear perception of all volume levels). The average of all ratings will be used as the final score.
Sound Files: Files will consist of talking normally (1.0, 1.1), talking loudly to whispering (5.0), instrumental piano (2.0), alarm sounds (4.0), squash ball sounds (4.1), and classroom background noise (3.0).

5. Frequency Range
Frequency range refers to the range of frequencies (low to high) that the processor can accurately handle. A wide frequency range allows for more natural audio reproduction, especially in the case of music and other environmental sounds.

Subjective Listening Test: Audio samples representing a variety of frequencies (such as music clips) will be played. Participants will focus on the clarity of both low and high frequencies, deciding whether all tones are easily discernible. Based on the performance of each processor design across multiple sound files, the frequency range will be rated on a scale of 1 (poor representation of frequencies) to 5 (excellent reproduction of all frequencies). The average rating for each sound sample will be used for the final score for the design.
Sound files: Instrumental Piano (2.0)

Decision Matrix
Following the creation of multiple signal processor designs, we will distribute and assess the synthesized output sound for each of our recorded sound files based on the rating scale described above. The sound files are intended to cover the range of sounds that are necessary to properly evaluate each criteria and that the typical user will encounter. As these results can be subjective, it is important to get a range of opinions on the sound output, and we will take the average of the results for the final scoring.

Each participant will then rank the option's output for associated sound files and we will place the total averages per sound file in the decision matrix below:

Design Option	Clarity	Word Intelligibility /Similarity to Input	Adaptability	Volume Range	Frequency Range	**Total**
Option 1	1: [1 to 5]	1: [1 to 5]	1: [1 to 5]	1: [1 to 5]	1: [1 to 5]	
	2: [1 to 5]	2: [1 to 5]	2: [1 to 5]	2: [1 to 5]	2: [1 to 5]	
	3: [1 to 5]	3: [1 to 5]	3: [1 to 5]	3: [1 to 5]	3: [1 to 5]	
	…	…	…	…	…	
Average	[1 to 5]	[1 to 5]	[1 to 5]	[1 to 5]	[1 to 5]	**[5 to 25]**

Note: For criteria that apply to multiple sound files, we will take the average from our rating of each of the inputs to get the final rating.

Sound Files (See deliverables for reference plots)

Speech Samples:

1.0 Male Speech (Male_Speech.wav): This sound file consists of a male speaker talking at a normal volume in an otherwise quiet environment.

1.1 Female Speech (Female_Speech.wav): This sound file consists of a female speaker talking at a normal volume in an otherwise quiet environment.

Music Samples:

2.0 Instrumental Piano (Instrumental_Piano_Sound.wav): This sound file consists of an instrumental piano, in an otherwise quiet environment.

Environmental Samples:

3.0 Classroom (Classroom.wav): This file consists of sounds from a lecture hall, consisting of multiple speakers, and other background noises (typing, chairs moving).

3.1 Street (Street.wav): This file consists of sounds taken from a busy street, consisting of people walking/talking, cars driving, the wind, and other environmental noises.

Impulse Samples:

4.0 Alarm Clock (Alarm_Clock.wav): This file consists of an alarm clock going off abruptly in an otherwise quiet environment.

4.1 Squash Ball Sound (Squash_Ball.wav) This sound file consists of a squash ball being hit in an echoey room using a racket with short pauses in between.

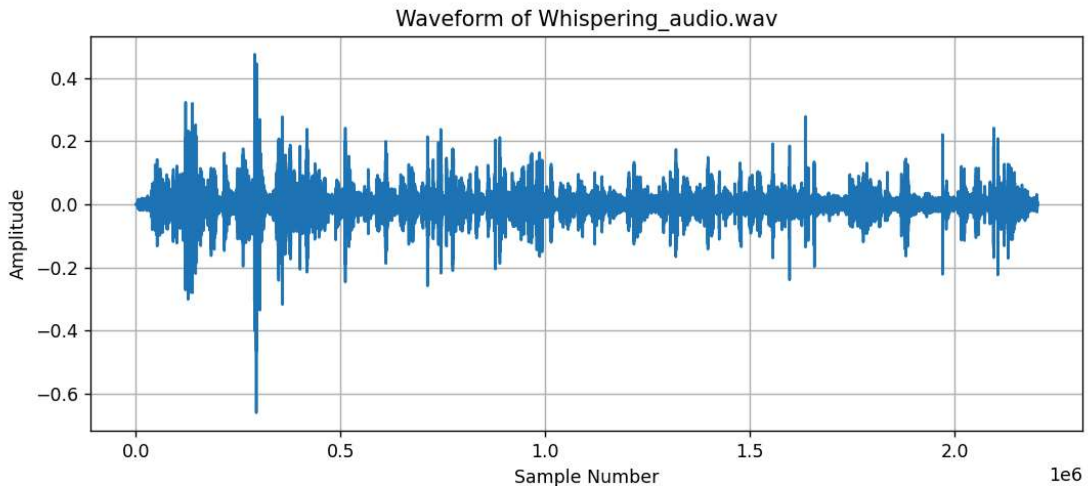
Changing Condition Samples:

5.0 Talking Normal To Whispering (Talking_Normal_To_Whispering.wav) This sound file includes two female speakers going from normal speaking to whispering, in an otherwise quiet environment.

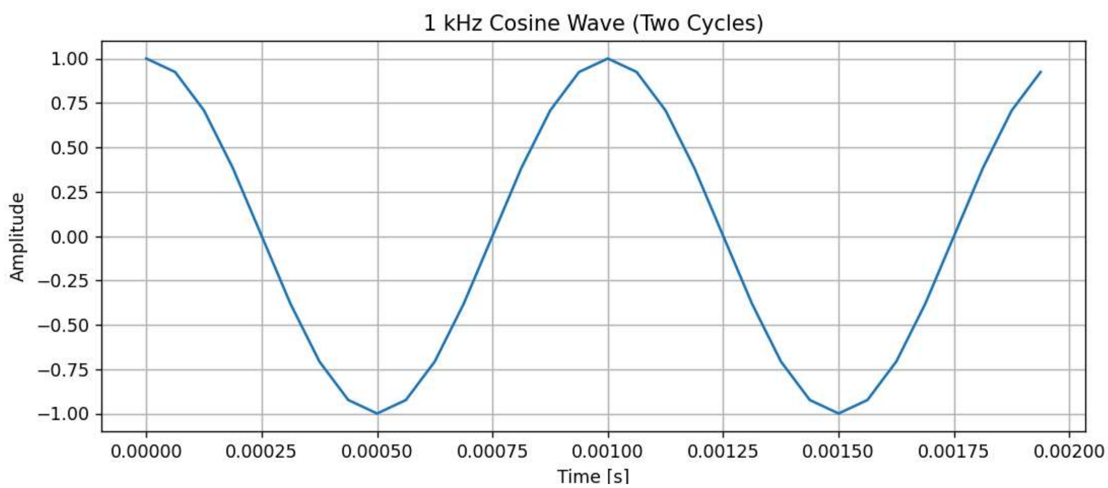
5.1 One Speaker to Multiple (One_Speaker_To_Multiple.wav) This sound file consists of one speaker talking at first, and then multiple people speaking.

Deliverables for Task 3

Plot 3.5



Plot 3.7



References

[1] F.-G. Zeng, S. Rebscher, W. Harrison, X. Sun, and H. Feng, “Cochlear implants: System design, integration, and evaluation,” IEEE Reviews in Biomedical Engineering, vol. 1, pp. 115–142, Jan. 2008. doi:10.1109/rbme.2008.2008250

[2] N. Deep, E. Dowling, D. Jethanamest, and M. Carlson, “Cochlear implantation: An overview,” Journal of Neurological Surgery Part B: Skull Base, vol. 80, no. 02, pp. 169–177, Sep. 2018. doi:10.1055/s-0038-1669411

[3] A. Snyder, “Making Music Enjoyable for Cochlear Implant Users,” NIH Record, Jul. 08, 2022. Available: https://nihrecord.nih.gov/2022/07/08/making-music-enjoyable-cochlear-implant-users

[4] R. Guski, “Psychological Methods for Evaluating Sound Quality and Assessing Acoustic Information,” Acta Acustica united with Acustica, vol. 83, Oct. 1997.