Fall 2018

Music Perception in Simulations of Cochlear Implant Listening

Elizabeth Mcnichols
Elizabeth.B.Mcnichols@Colorado.EDU

Music Perception in Simulations of Cochlear Implant Listening

Elizabeth B. McNichols

October, 31, 2018

Committee Members:

Kathryn Arehart: Speech, Language, and Hearing Sciences

Naomi Croghan: Speech, Language, and Hearing Sciences

Fernando Rosario: Civil, Environmental, and Architectural Engineering

Ann Schmiesing: German Studies

Abstract

Cochlear implant (CI) processing has been optimized for speech perception, but music perception has been a secondary consideration. A proposed signal processing strategy called focused stimulation may help with music perception for cochlear implant users. This strategy aims to improve spectral resolution (compared to previous signal processing strategies) by reducing the amount of current spread that occurs in the CI electrode array. In the following experiment, 14 normal hearing young adults listened and rated the sound quality of music samples that were processed to simulate a CI with various amounts of simulated electrical spread. Ratings were performed using MUltiple Stimulus Hidden Reference and Anchor (MUSHRA) protocol. Input resolution was manipulated through spectral smearing and acoustic differences in genre. It was found that with more electrical spread, participants had difficulty hearing changes in spectral resolution. In addition, the effects differed across musical genres. The results show that minimizing electrical spread improves spectral resolution in normal hearing participants, but these effects need to be tested on CI users.

Literature Review

Music is magic. It gives us the ability to time travel. Like when I hear "Baby" by Justin Bieber I am instantly transported back to awkward middle school dances (Bieber, 2009). Music has the power to unite. Like when "Sweet Caroline" comes on at a football stadium and there is a resounding "Sweet Caroline, ba-ba-ba" from fans rooting for both teams (Diamond, 1969). Music can give us an identity and culture. Like when we see tie-dye-clad Dead Heads invade Boulder every summer for the Dead and Company concert.

Imagine a world without music. A world where the intricate guitar solos of Jimi Hendrix sound like a series of beeps. Where you could barely differentiate between the deep wailing of a cello versus a playful flute. For individuals with cochlear implants (CIs) this is the reality of listening to music. For these CI users, the magic of music is diminished because of the capability of their devices. While some may believe music is secondary to speech in a person's well-being, the literature shows music is essential to life (North & Hargreaves, 2003).

Researchers recently conducted interviews of young people (aged 15-25) to examine the effect music had on their life (Papinczak et al, 2015). The researchers identified four major themes. (1) It was found that music listening built relationships through sharing music with friends and attending concerts. (2) Music could also modify cognition through aiding concentration, evoking positive or negative memories associated with the music, or as a tool to solve problems by listening to the messages in lyrics. (3) The research found music was used to modify emotion through distractions and altering alertness. For instance, subjects reported playing upbeat music to increase energy levels or listening to relaxing music before going to bed. (4) Participants reported using music as a way to intensify emotions. Participants would listen to certain songs when they are sad to fully embrace their emotions and initiate the healing process.

Youth with CIs do not exhibit the same relational and emotional benefits from CIs due to the poor music quality.

These finding are not exclusive to the young adult population. A study conducted on elderly people found that music promotes a strong sense of self-identity through reminiscing (Dasa, 2018). In the same study, music was also found to help subjects cope with loss and change. These positive outcomes led to an increase in socialization, self-esteem, and general life satisfaction. The positive mood benefits were confirmed by another study which found that when music was used as therapy for depression, there was a significant reduction of the symptoms of depression (Zhao et al, 2016).

It is clear music is crucial for many aspects of the human experience. Individuals with hearing aids and CIs do not experience music in the way normal hearing individuals do which effects their quality of life (Dritsakis et al, 2017). Part of the reason for this discrepancy is due to the difference in how a healthy ear and CI process sound.

In a functioning auditory system, sound from our surroundings is funneled into our external auditory meatus. The pressure from the sound waves then moves the tympanic membrane, and in turn, vibrates the middle ear ossicles. The footplate of the stapes then pushes the oval window of the cochlea. Because the cochlea is a fluid filled structure, the mechanical information from the middle ear ossicles is transferred to the basilar membrane in the form of a traveling wave. The up and down movement of the traveling wave causes the hair cells in the Organ of Corti to bend which in turn causes the auditory nerve to fire. Importantly, the physical properties of the basilar membrane cause it to be a frequency analyzer. That is, the stiff base responds best to high frequencies and the less stiff (flaccid) apex responds best to low frequencies. This frequency by place map is called tonotopic organization and is one of the important ways that complex sounds in music are pulled apart and analyzed by the auditory system.

A healthy ear has two mechanisms that aid in sound sensitivity and frequency resolution: active and passive mechanisms. When living cochlea are studied, the actions of the active mechanism are evident in the system's frequency selectivity and sensitivity (Gold, 1948). The main structure involved in the active mechanism are the outer hair cells (OHC). Motility of the OHC cause them to expand and contract to amplify and sharpen the traveling wave, causing frequencies to be more easily detected and resolved (that is, to separate one frequency from another happening at the same time). This added movement increases frequency resolution through improved selectivity and specificity (NIDCD, 2015). The passive mechanism has been studied in postmortem cochlea and consists of the physical properties (mass and stiffness) of the basilar membrane. This passive mechanism provides some frequency selectivity, but does not provide nearly as much as is seen when compared to a healthy cochlea (Gold, 1948).

With the help of both mechanisms, the cochlea acts as a series of band pass filters which allow us to pull apart and resolve the many sine waves making up the complex sounds in speech and music. This precise spectral resolution is why the healthy cochlea is often called a "Fourier analyzer" (Plack, 2013). Figure 1 (Plack, 2013) provides a visual representation of the auditory filters. The horizontal axis represents frequency and the vertical axis represents the intensity of the filters. This is a conceptual schema; in reality, there are more filters that are spaced closer together.

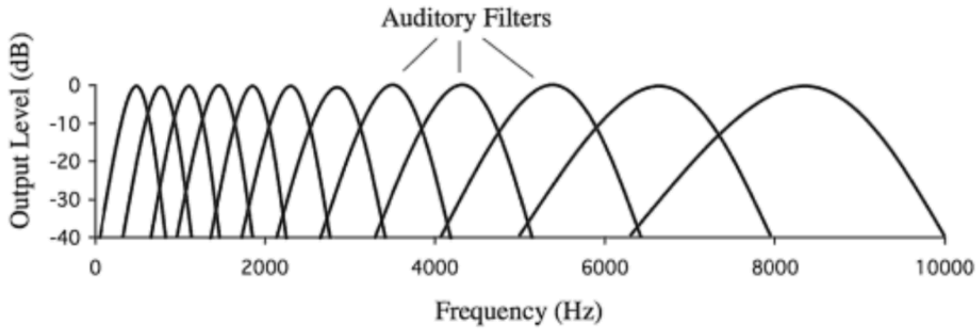

Figure 1 is a conceptual drawing of auditory filters. In reality there are more auditory filters and they are closer together. On the horizontal axis is frequency and sound intensity is shown on the vertical axis. (Plack, 2013).

In comparison, CIs attempt to process sound similar to a functioning auditory system, but by using technology. The prosthetic device consists of a microphone, signal processor, and electrode array. Sound is first picked up by the microphone and is converted to an electrical signal that is processed in a way that attempts to mimic the frequency analysis that is done by a normally functioning cochlea. That is, the incoming sound (e.g,, a musical piece) is put through a bank of band pass filters. The components of the incoming sound that fall within a particular band will be mapped to a particular electrode that then directly stimulates the auditory nerve. For example, a low note of a melody will cause the most stimulation of electrodes that are near the apex of the cochlea. A high frequency note will cause stimulation of electrodes that are closer to the base of the cochlea. The electrodes then stimulate the auditory nerve, which transmits this neural information to the brain.

One of the main differences in hearing with a healthy auditory system and hearing with a CI is that the filtering and spectral resolution provided by the healthy auditory system is more detailed and precise. The reason that CIs have worse spectral resolution is because of the effective number of filters is smaller. Commercial CIs have up to 22 channels on their electrode arrays. In the past, research has shown that CI users only receive benefit in hearing from approximately eight electrode channels, especially when listening to speech in background noise

(Fishman et al, 1997; Frisen et al, 2001; Garnham et al, 2002). However, with improvements in technology and surgical techniques, CI users may receive benefit from all the electrode channels (Croghan et al, 2017). Regardless, even if all the electrode channels are useful, the spectral resolution is still not as precise as that of a normal ear.

The difference in frequency analysis between a healthy cochlea and CI has implications for hearing speech and music. CI recipients generally have good speech recognition when speech is presented in quiet (Gifford et al, 2008). However, when environmental noise interferes with the speech signal, speech recognition is diminished (Brant et al, 2018) since good spectral resolution is even more important for understanding speech in noise (Lorenzi et al, 2016).

As with noisy speech, music perception also requires very precise spectral resolution (Oxenham, 2008). Because CIs lack this precise spectral resolution, music perception with CIs is not nearly as good as with normal hearing. Consider, for example, the results of a study comparing musical enjoyment for people who have a CI in one ear (due to single-sided deafness) and normal hearing in their other ear (Landsburger, 2017). Music heard through the CI ear was rated with significantly less musical enjoyment, even if the listeners had significant experience with CIs (Landsburger, 2017). This discrepancy is in part because of the difference in spectral resolution between a healthy auditory system and a CI.

In this thesis, spectral resolution can be examined in terms of input and output spectral resolution. Input resolution consists of the acoustic properties of the sounds that go into the auditory system. Output resolution consists of the spectral degradation which occurs from constraints in the device.

For this thesis, input resolution is varied by spectral smearing and acoustic variation from music genres. As described previously, spectral resolution is integral to perceiving differences in frequency such as in speech and music perception. Researchers have utilized this relationship to build stimuli for testing aspects of CIs. For instance, in Litvak et al (2007), the researchers tested both normal hearing individuals and CI users to find the smallest spectral contrast between two stimuli when the number of effective channels was manipulated. Individuals with normal hearing listened to the stimuli after it had been processed in a way which simulates CI processing. Similarly, Smith et al (2013), used a test with different stimuli (but that still varied in spectral resolution) which has been correlated to with speech understanding both in quiet and in noise. The purpose of this study was to test outcomes of CIs with different numbers of effective channels (Henry et al, 2005; Won et al, 2007). By varying the spectral resolution of the stimuli, researchers were able to determine the effectiveness of CI sound processing strategies based upon listeners' ability to detect spectral degradation in the stimuli.

Another factor to consider regarding the input signal is its unique acoustic properties. Music and speech differ in their acoustic properties. However, amplification devices and CIs have been designed to process sound in a way that has been optimized for speech perception. Generally, music has more variation than speech in spectral components. This is due to to the various instruments playing at different tempos and rhythms (Croghan, 2013). The acoustic differences between speech and music pose issues for music perception. Music perception depends on discerning and resolving small differences in frequency. For example, pitch perception depends on being able to resolve the sine waves that are contained in a complex sound played by a musical instrument. Most normal hearing individuals can detect half-step changes in frequency whereas CI users range from being unabl to detect a difference in frequency until the distance is anywhere from one to eight half steps (Kang et al, 2009). This inability to perceive smaller changes in pitch is one contributing factor with difficulty in perceiving pitch and melody (Limb, 2014).

The output resolution in this thesis is examining the number of useful channels in CI processing strategies. Commercial CIs utilize a processing strategy called monopolar stimulation. In the monopolar system, the active electrode is in the cochlea and one of the grounding electrodes is outside of the cochlea. Because these two electrodes are far apart, the electrical current spreads further on the electrode array than intended and stimulates a larger neural population (Figure 2; courtesy of Naomi Croghan, Cochlear Ltd). The current spread between channels is referred to as channel interaction. With more channel interaction, the frequency fidelity of the signal is compromised resulting in diminished clarity. If one were to imagine the tonotopically organized cochlea as a piano, current spread would be like playing a song with one's forearm; the desired key (frequency) would be played (stimulated), but so would the surrounding keys. This compromised frequency selectivity is one component contributing to poor music perception in CIs.

One cochlear implant company -- Cochlear Ltd. -- has been investigating an alternative method of stimulation called focused stimulation. In focused stimulation, the desired electrode is stimulated with a positive current. Simultaneously, the surrounding electrodes receive a negative current to cancel the electrical spread. Thus, there is less channel interaction resulting in better frequency selectivity (Figure 2; courtesy of Naomi Croghan, Cochlear Ltd). Extending the piano analogy, in focused stimulation, one is now playing with one's fingers. With focused stimulation, less unintended information is conveyed, thus improving frequency selectivity and music perception.

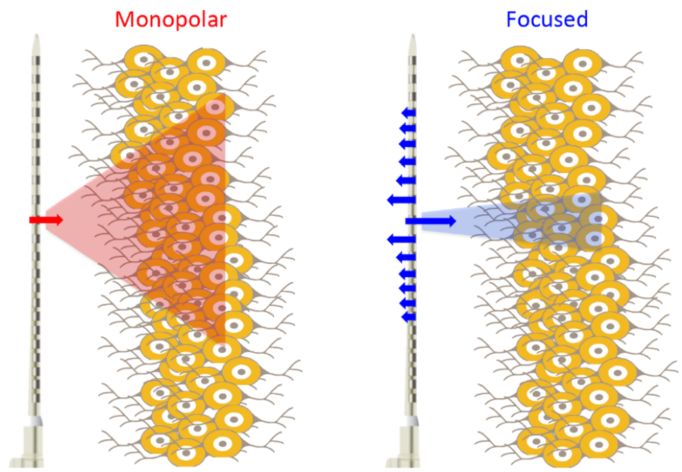

Figure 2 shows the electrode array and stimulation pattern for monopolar and focused stimulation. The left side shows electrical spread in monopolar stimulation and the wide neural region it stimulates. The right side shows the electrical spread in focused stimulation. The blue arrow pointing right indicates the positive current going towards the nerve cells and the arrows pointing left represent the negative current being sent to adjacent channels to minimize the channel interaction and stimulate a more narrow neural region. Courtesy of Cochlear Ltd.

Cochlear implant manufacturers often want to assess how a new processing strategy compares to an existing one. These assessments typically involve research studies that compare music perception in two ways:

1. Using an experimental implant system with a CI listener
2. Using CI simulations to test individuals with normal hearing

There have been several studies which have used either one or both of these methods to test signal processing strategies in CI.

For instance, Roy et al (2015) compared the effectiveness in a new processing strategy compared to an existing strategy in a commercial CI system. The new processing strategy was designed to provide better frequency resolution which has been shown to be critical for good music perception. In the study, participants listened to music samples that had been high pass filtered to remove different amounts of low frequency information. The stimuli were filtered in order to test participants' sensitivity to low frequency changes when listening to both of the processing strategies. Participants were then asked to rate the music quality of the samples using a method called MUltiple Stimuli with Hidden Reference and Anchor (MUSHRA). Participants were asked to compare the degraded samples to the unprocessed version (reference). Within each set of samples there was one identical to the reference and one that was heavily degraded (anchor). These hidden stimuli helped to encourage listeners to use the whole scale for ratings and tease out differences between the different processing conditions. The CI participants listened to the samples twice, once with their devices set to the new processing strategy that provides better spectral resolution. The device settings were then changed to the older processing strategy which participants wore for two months to become adapted. The quality ratings were then repeated with the older strategy. Individuals with normal hearing completed the same test without any processing as a control and their results were compared to the CI results. It was found that the results of the newer CI processing strategy more closely resembled the results obtained from the normal hearing control subjects. This finding means that the processing strategy that better encodes the fine structure information improves musical sound quality.

The electrode array in CIs often do not reach the apical end of the cochlea and therefore do not stimulate the low-frequency regions. Some companies are now offering an electrode that can be inserted deeper into the cochlea in the hopes to utilize the low-frequency regions that are typically neglected. A study by Roy et al (2016) investigated the effect of deeper electrode insertion on music quality in CI users. The participants were CI users with different electrode array lengths along with normal hearing listeners. The participants listened to music stimuli that had been filtered so that there were different amounts of low frequency information. This was used to test participants' sensitivity to low frequency information. Participants listened to the samples and then rated them to how similar they sounded to the unfiltered reference. Following the MUSHRA protocol, there was one sample identical to the reference and one highly degraded sample to act as the anchor within each set of stimuli. The data from the CI listeners were then compared to the data from the listeners with normal hearing. It was found that the individuals with deeper inserted electrodes had MUSHRA ratings more similar to listeners with normal hearing, suggesting apical stimulation improves low-frequency perception. This may transfer to better music perception for CI users.

Smith et al (2013) compared music perception in CI users for the commercially used monopolar stimulation and the experimental focused stimulation. The nine participants had devices that could be programmed to have the same current interaction as either monopolar or focused stimulation. With each of the processing strategies, the participants performed a task which tested their perception of spectral information. The ability to preform these tasks has been found to indicate better speech understanding in complex listening situations. Based off of the data collected from the CI users, it was found that participants utilizing focused stimulation were better able to perform the tasks than with monopolar stimulation. The results indicate that focused stimulation can improve spectral resolution, thus increasing the number of useful channels through minimizing the channel interaction. Improved spectral resolution has implications in improving sound perception in complex or noisy conditions. While the findings from Smith et al (2013) are integral to understanding the potential of focused stimulation, more research needs to be conducted on the efficacy of focused stimulation in other listening conditions such as music. This thesis aims to investigate the effect of electrical spread on music perception by testing individuals with normal hearing who listen to music that is simulated to sound like CIs.

The main effect that was tested in the study was electrical spread within the electrode array. This was simulated using a process called vocoding. Based upon the Smith et al (2013) study, with less current interaction there is improved frequency selectivity, thus improving sound perception. Individuals who listen to music samples with less electrical spread should be better at discerning spectral information in music due to better frequency resolution. Considering this, a primary experimental question is: ***how do various amounts of simulated electrical spread affect music quality perception in simulations of CI?***

The next factor to be considered in this experiment is spectral smearing. This was simulated through a signal processing technique called cepstral analysis. Spectral smearing has been used as a tool to indicate listeners' ability to detect spectral degradation with the different amounts of electrical spread (Litvak et al, 2007; Smith et al, 2013). This begs the experimental question: ***how does various amounts of spectral smearing, when combined with different amounts of electrical spread, affect music quality perception?***

The final factor of interest is the effect of acoustical differences on music quality ratings. We know that CIs work well with speech in quiet, but their processing techniques do not transfer to the unique acoustics of music. Even within music there is variation in the rhythm, melody and tempo based off of the notes and instruments in a song. Because of the wide range of music CI users will listen to, the question arises: ***how do different genres of music affect music quality perception when different amounts of electrical spread and spectral resolution are present?***

The following study aims to address these three questions using simulations of CI processing. That is, normal hearing individuals listened to songs from different genres that were processed to simulate different amounts of electrical spread and spectral resolution.

	Factor	Significance	Impact on music	How represented in the simulation	Experimental question
Output Resolution	Electrical Spread	When an electrode on the array is stimulated, the current spreads to the adjacent channels. By changing the processing strategy, the channel interaction can be mitigated.	With less electrical spread, a smaller area of the basilar membrane will be stimulated, and frequency selectivity will be improved. This translates to possible better music perception.	*Vocoding* • Unprocessed (least spread) • -40 dB/channel • -12 dB/channel • -6 dB/channel (most spread)	How do various amounts of simulated electrical spread affect music quality perception in simulations of CI?
Input Resolution	Spectral Resolution	Degrading spectral resolution causes sound samples to become unclear. By manipulating resolution before applying a processing strategy, subjects' ability to detect spectral changes can be tested.	Subjects' ability to detect spectral degradation in different electrical spread simulation will indicate improved music quality perception.	*Smearing Factor* • Reference (least degraded) • 0.7 • 0.5 • 0.4 • 0.3 • 0.2 • 0.1 (Anchor, most degraded)	How does various amounts of spectral resolution, when combined with different amounts of electrical spread, affect music quality perception?
	Acoustic Differences	Each instrument has unique acoustic properties. Perception may vary depending on the type and number of instruments.	The music genre may have an impact on the quality ratings with different amounts of spread and smearing.	*Genre* • Rock • Rock Simple • Folk • Folk Simple • Flute • Classical	How do different genres of music affect music quality perception when different amounts of electrical spread and spectral resolution are present?

Table 1 outlines the 3 main factors being studied in this experiment, the impact those factors have in music, how they are being simulated in the experiment, and their corresponding research question.

Methods

Subjects

Fourteen young adults (19-29 years old; average age of 24) with normal hearing participated in this experiment. Hearing status was determined by a hearing screening. Listeners sat in a sound attenuated booth and listened to pure-tones at octaves between 250-8000 Hz. The tones were presented using an Grason-Stadler AudioStar Pro audiometer over TDH-49 headphones. Listeners were required to respond at 20 dB HL at each octave in order to qualify for the study.

Of the fourteen participants, four were self-proclaimed musicians and one had received some audio-recording training. Listeners reported listening to recorded music 0-40 hours per week (average of 12 hours). The research participants consisted of 10 females and 4 males. Listeners were recruited by methods such as flyers that were hung in heavily trafficked areas of campus and postings to E-bulletin boards. The advertisements and dissemination methods were approved by the University of Colorado at Boulder Institutional Review Board (IRB).

Stimuli and Signal Processing

Participants listened to six different music samples throughout the study (Table 2). The samples consisted of rock, folk, orchestral, and a single flute. Both the rock and folk had two versions: the original and a version with some of the background instruments removed for simplicity.

Genre	Song Title	Artist
Rock Complex	“Schoolboy Fascination”	Al James
Rock Simple	“Schoolboy Fascination”	Al James
Folk Complex	“Rachel”	Southern Sirens
Folk Simple	“Rachel”	Southern Sirens
Orchestra	“Ode to Joy”	Beethoven
Flute		

Table 2 describes the different songs used as stimuli for the experiment. Both Rock and Folk have two versions: the original (complex) or a version in which background instruments were removed for simplicity.

The six music samples for this study were processed in two steps. The first step effected the input resolution of the sound. This step models spectral smearing in the CI which has been connected to CI users’ performance in complex listening situations. The music samples were processed using cepstral analysis to mimic this effect.

For the ceptral analysis, a Fast Fourier Transform (FFT) was performed (Litvak et al., 2007). The result was the frequency spectrum of the input. The absolute value and logarithm of the positive frequencies was performed. At this point, some of the spectral variation was removed systematically (Figure 3). The amount of information removed was dependent on the smearing factor. For instance, if 60% of the spectral variation is removed and 40% remain, the smearing factor would be 0.4. If there was no spectral variation taken out, the smearing factor would be 1. After the designated amount of spectral variation was removed, the process was reversed. The output was the original sound sample missing the desired amount of spectral information.

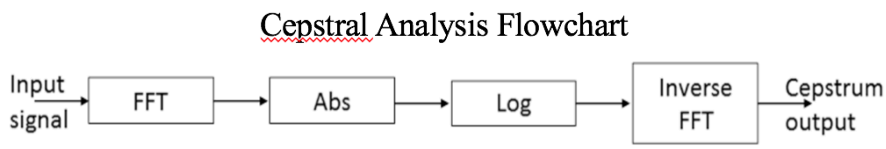

Figure 3 explains the processing steps in cepstral analysis. FFT stands for Fast Fourier Transform, Abs is an abbreviation for the absolute value, and log means the logarithm of the values was taken.

The frequency spectrum in Figure 4 (courtesy of Naomi Croghan, Cochlear Ltd.) shows the effects of cepstral analysis on a sound. The information in blue is the original sound sample whereas the information in red represents the processed sound. By comparing the two, one can see the processed sample has more gradual curves as opposed to the spikes in the unprocessed sample. This is a visual representation of the loss of spectral variation from a cepstral analysis.

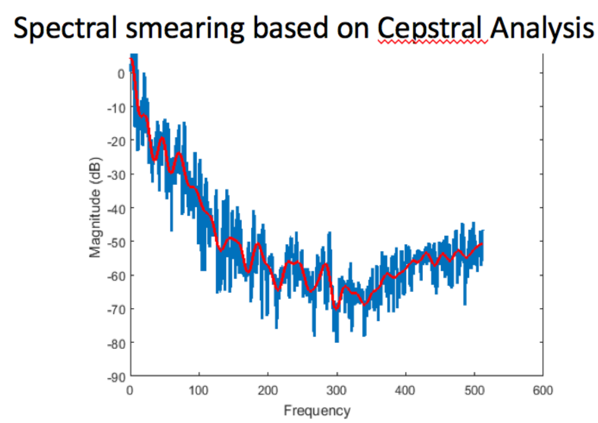

Figure 4 is a spectrum of two sound samples overlaid on each other. The blue line indicates an unprocessed stimulus with many clear peaks and valleys. The red line represents the same stimulus but with spectral information removed as evidenced by the more gradual peaks and valleys. Courtesy of Cochlear Ltd.

If one were to think of a music sample as a picture, spectral smearing can be compared to pixilation (Figure 5). The more smearing, the less clear the picture is because the spectral information is being removed. This processing step makes the sound sample sound “muddied”.

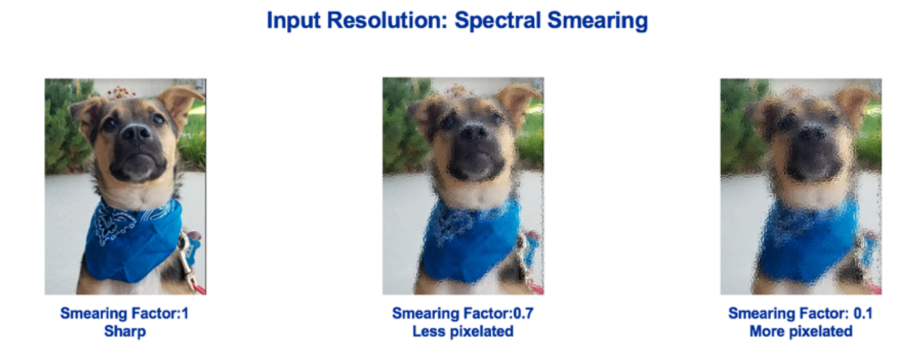

Figure 5 demonstrates the concept of the effect of spectral smearing on spectral resolution through the pixilation of a photo. In this analogy, the photo is the sample and the pixilation is the smearing factor. With more smearing, the photo becomes less clear. This translates to spectral smearing of the music samples. With different amounts of spectral smearing, the sample sounds less clear because of the missing spectral information.

The second type of processing simulates various processing strategies in cochlear implants. The samples with spectral smearing were processed to mimic the processing strategies of a CI through a process called vocoding (Figure 6). This step impacts the output resolution (simulated current spread) of the music. A music sample was broken into various frequency bands that span the frequency range of the human ear via band pass filters. Then the envelope of each frequency band was extracted. In this simulation, a decay function was applied to the envelopes. This mimicked the various processing strategies available through simulating the electrical spread within the device. The more decay, the less current spread there is, and the clearer the music samples will be. Subjects listened to music samples that were only smeared (with no vocoding), -40dB/channel, -12 dB/channel, and -6 dB/channel. The vocoding values indicate how much weight the vocoding output has on the surrounding signal. The higher the value, the less weight, and the clearer the sample sounds. For instance, -40 dB/channel is less “blurred” than -6 dB/channel because -40 dB/channel puts less emphasis on the surrounding frequency outputs. Thus, -40 dB/channel has better frequency resolution and sounds clearer than the -6 dB/channel. The envelope was then imposed over a carrier noise. Through this simulation of signal processing, normal hearing listeners can experience music the way the CI processes it.

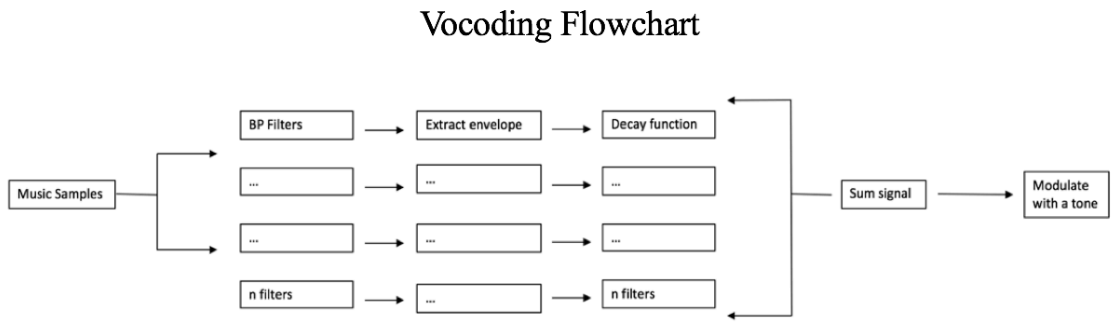

Figure 6 demonstrates the signal processing steps to achieve vocoding. These steps attempt to simulate the way a CI processes and encodes sounds. A series of band pass filters are applied to a music sample to break the sample into frequency bands. The envelope of the sample is extracted and then a decay function is applied in order to simulate different amounts of electrical spread. Each frequency band is then added together and modulated with a tone.

For this experiment, the stimuli that were unprocessed are relevant because they reflect what subjects with CI listened to as part of a larger study being conducted by Cochlear Ltd on CI users. The -40 dB/channel represents a best case scenario for frequency resolution. -12 dB/ channel simulates the electrical spread seen in focused stimulation. Finally, -6 dB/ channel represents the electrical spread seen in monopolar stimulation. These conditions give us the ability to compare the efficacy of various processing strategies and the effects of simulated electrical spread within typical hearing individuals. Table 3 summarizes the simulated electrical spread conditions.

Simulated Electrical Spread Conditions

Simulated Electrical Spread
Unprocessed (UNP)
Neg40
Neg12
Neg6

Least Electrical Spread
More Clear

More Electrical Spread
Less Clear

Table 3 demonstrates the degrees of simulated electrical spread in the experiment. With an increase in electrical spread, the samples become less clear.

Testing

Each participant took part in a total of three, one-hour sessions. During the first session, listeners began with a hearing screening and a survey regarding their musical backgrounds. The survey was disseminated utilizing the University of Colorado Qualtrics survey software. Participants filled out the survey in lab so they could ask any clarifying questions about the prompts. The survey consisted of three subsections. The first asked if the subject had any previous musical training, what type of training they had, and how long did they train. The second section of the survey asked about any past audio recording experience, education, or training. Finally, the survey asked about music listening habits of the individual (how long they listen/day, what music genre, etc.). Previous musical experience was not grounds for disqualification from the study, but these data told us more about the listeners we had recruited.

Subjects then rated the quality of music samples according to the MUSHRA protocol. The stimuli were played through Matlab and were routed through a GSI Audiostar Pro audiometer. The stimuli were presented using the audiometer's insert headphones and was played at 65 dB SPL.

Each testing session started with a training phase in which the participants became familiarized with the range of music samples. The participants were asked to click on each of the boxes as seen in Figure 7 below. Through this step, listeners became acquainted with the reference and the various degrees of degradation they should expect from the subsequent samples. Once subjects had familiarized themselves with the range of the stimuli, they clicked "proceed to evaluation" to start the evaluation phase of testing.

MUSHRA Familiarization Task

Figure 7 is the testing interface participants used for the familiarization step of testing. Each "experiment" is a different music genre. The reference is each music sample without any spectral smearing. Each of the test buttons per experiment are the same music sample with a different smearing factor. All the stimuli in this task has the same simulated electrical spread.

In the evaluation phase, participants listened to the reference and then rated the subsequent music samples to how closely they sounded to the reference (Figure 8). As according to MUSHRA protocol, the reference is hidden in the stimuli set. Participants were told this and instructed to rate at least one of the samples at 100 because one was a complete match to the reference.

MUSHRA Testing

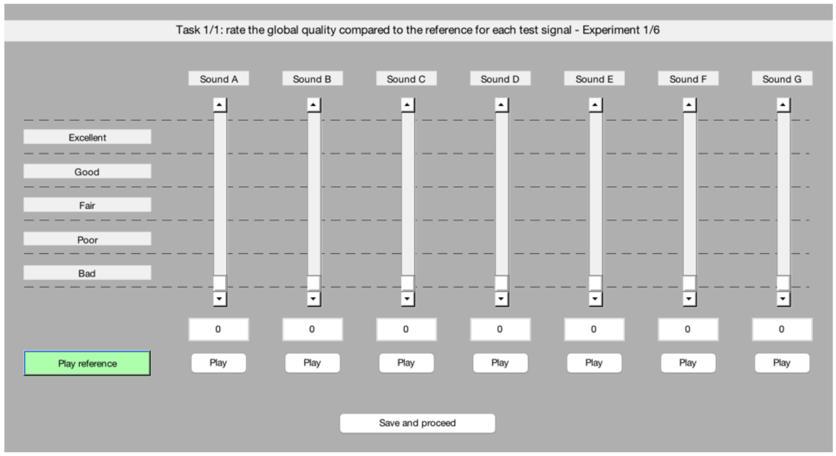

Figure 8 is the testing interface for the experiment. The reference is a sample with no spectral smearing. The sound samples A-G are all the same genre with the same simulated electrical spread, but all have differing amounts of smearing ranging from none (hidden reference) to heavily degraded (anchor). Participants listened to the reference and then gave the samples A-G a score between 0-100 on how similar the sample sounded to the reference.

This procedure was repeated for each vocoding condition over the three visits. During the first visit, subjects listened to the simulated electrical spread condition in order of least spread (UNP) to most spread (neg6). The purpose of this session was to familiarize the participants with the task and the stimuli. In the following two sessions the simulated electrical spread conditions were randomized. Because of this discrepancy, only the visits where the vocoded conditions were randomized were used in data analysis.

Results and Discussion

Test/Retest Reliability

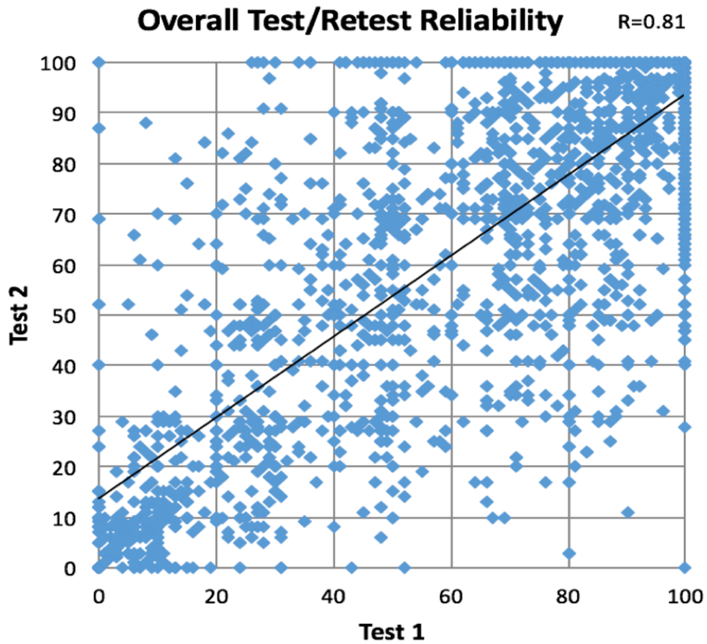

Figure 9 shows the test/retest reliability of participant's rating between their first testing session (horizontal axis) and second testing session (vertical axis). R=0.81.

Figure 9 shows the test/retest reliability of participants on the two sessions with randomized simulated electrical spread conditions. The graph shows the relationship between the quality rating for the first session (horizontal axis) and the quality rating for the second session (vertical axis). The Pearson Correlation Coefficient was 0.81. This positive correlation was statistically significant (p<L0.01) which means that the test-retest were closely related to each other.

Main effects

The purpose of this experiment is to determine the effects of simulated electrical spread on music quality perception when environmental factors (spectral resolution represented by smearing and acoustic variation represented through genre) were varied in simulations of CI listening.

Through this experiment, it was found that electrical spread had a significant effect on music quality perception. Figure 10 shows the overall ratings for electrical spread averaged across participants, smearing factors, and genre. A Repeated Measures Analysis of Variance (RM ANOVA) showed that the effect of simulated current spread was significant. As there is more simulated electrical spread, quality ratings increase. While that may seem counterintuitive, the task called for people to rate the processed samples compared to the reference. The higher scores occurred because individuals had difficulty discerning differences in spectral resolution and rated more of the samples as sounding similar to the reference. Post-hoc pairwise comparisons with Bonferonni corections also showed that the UNP condition was significantly different than the other spread conditions and that the neg40 condition was significantly different from neg6 (Table 4). The overall trends seen in Figure 10 demonstrate that with less simulated electrical spread, the perceived quality of music improves.

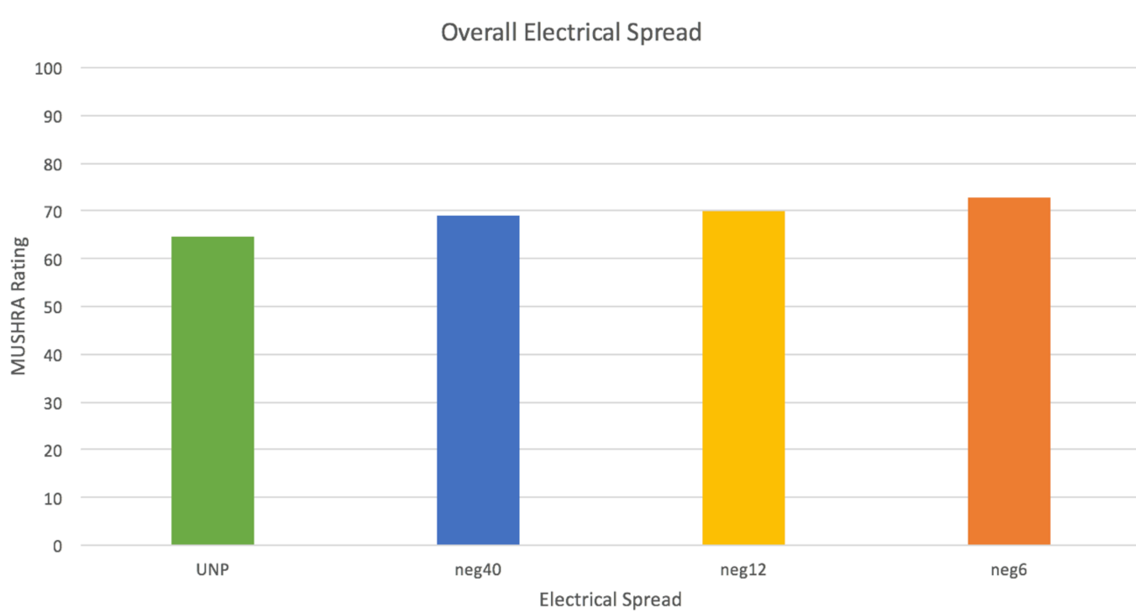

Figure 10 shows the main effect of electrical spread on MUSHRA ratings when averaged across participants, spectral smearing, and genre. Electrical spread is on the horizontal axis and MUSHRA ratings are on the vertical axis.

Another factor being examined was spectral smearing. In Figure 11, smearing was examined when averaged across participants, electrical spread, and genre. Generally, quality scores decreased with an increase in smearing factor because listeners heard the spectral degradation and gave them lower ratings. It was found that there was a significant difference in the quality ratings between each smearing factor except between 0.7 and 0.5 (Table 2). These findings signify that participants were able to detect spectral degradation in different music samples. With an increase in spectral smearing there is also a decrease in spectral resolution. These findings are consistent with the idea that good spectral resolution is essential for music perception (Oxenham, 2008).

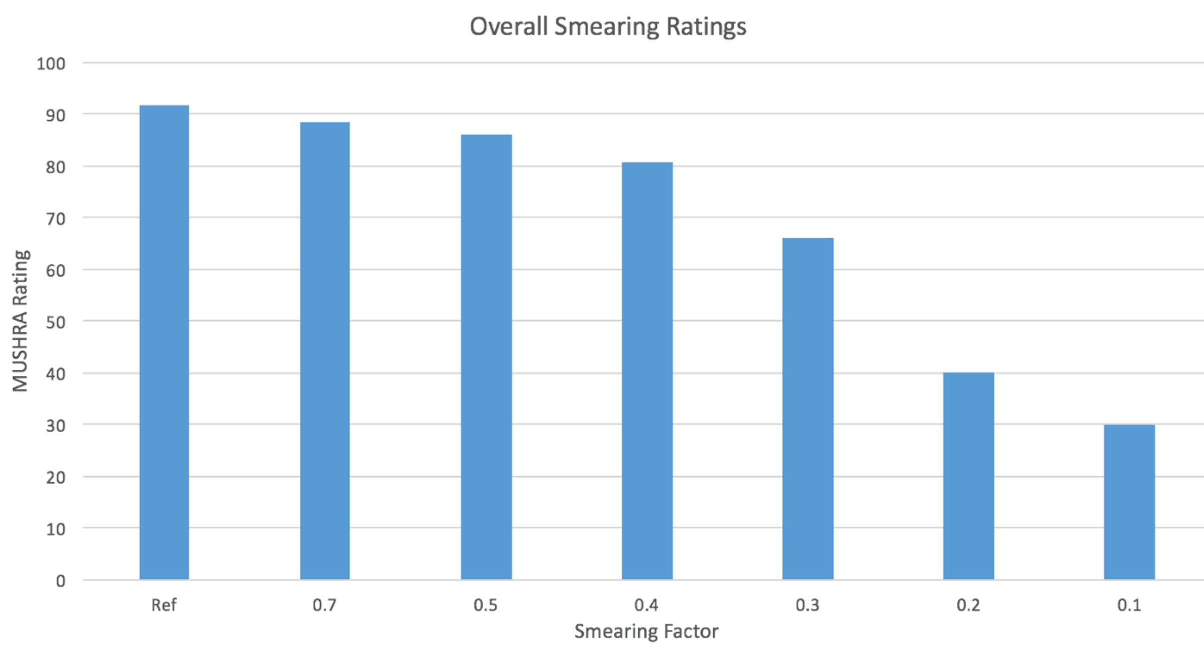

Figure 11 shows the main effect of spectral smearing when averaged across participants, simulated electrical spread, and genre. The smearing factor is on the horizontal axis and the MUSHRA ratings are on the vertical axis.

Genre was the final factor that was examined in this study. Figure 12 depicts the quality ratings for the six genres when averaged across participants, simulated electrical spread, and spectral smearing. Again, the main factor of genre was significant, meaning that the quality ratings differed for different genres. Through post-hoc pair-wise comparisons it was found that the Folk Complex condition was significantly different from all other genres except Folk Simple. In addition, Folk Simple was significantly different than all other genre besides Folk Complex and the Flute. These findings signify that music genre has an effect on music perception in simulations of CI.

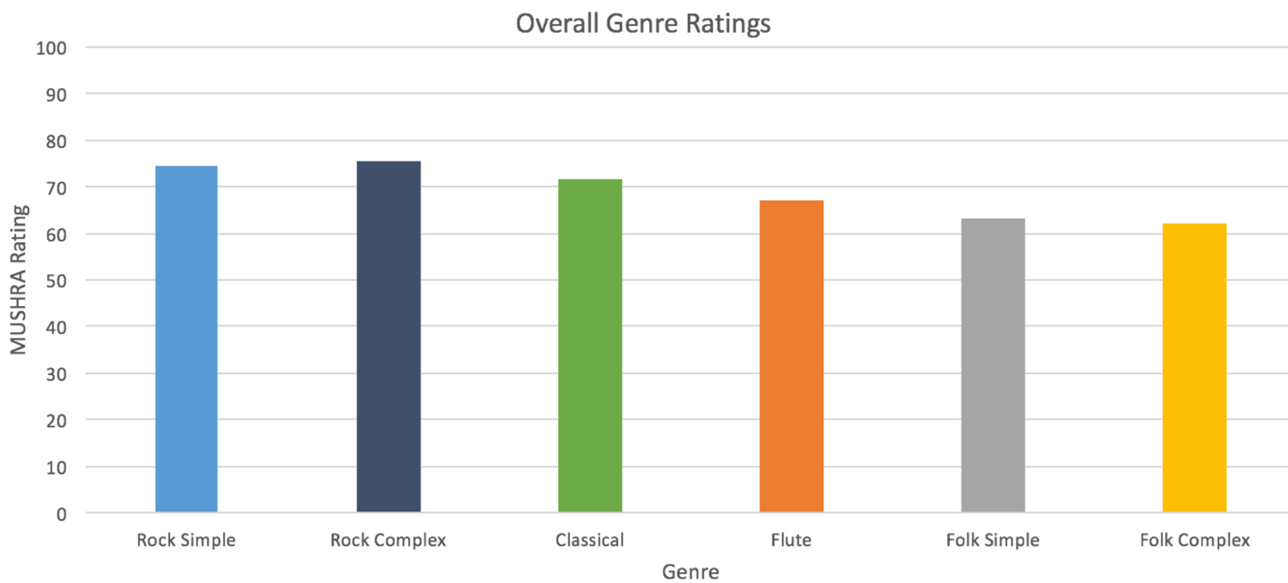

Figure 12 shows the main effects of genre when averaged across participants, simulated electrical spread, and spectral smearing. Genre is on the horizontal axis and the corresponding MUSHRA rating is on the vertical axis.

Interactions

Extending beyond the main effects, there were significant findings in the interactions of factors. For instance, Figure 13 demonstrates the interaction of electrical spread and smearing. On the left graph are the quality ratings averaged over genre and participants. On the horizontal axis are the electrical spread conditions denoted in different colors. Each bar within an electrical spread category are the smearing factors as labeled. The right graph denotes the difference between the reference and 0.1 (anchor) quality scores for each electrical spread condition. On the horizontal axis are the difference in quality scores and the vertical axis are the electrical spread conditions.

Figure 13 (left) Quality rating scores with spread condition on the horizontal axis and quality rating on the vertical axis. The 7 smearing conditions are represented by the bars in each electrical spread condition.
(right) difference between the reference and 0.1 (anchor) ratings for each electrical spread condition. Electrical spread is on the vertical axis and the quality rating is on the horizontal axis.

Two trends emerge from these graphs. (1) Generally, quality scores decrease as there is more smearing in each electrical spread condition. This is seen by the negative slope in the bars on the left graph. (2) With more electrical spread, the difference between the reference and anchor decreases. The UNP difference score is about 75 whereas the neg6 condition is closer to 50. One explanation is that listeners had more difficulty hearing differences in spectral resolution when there was more simulated electrical spread. Therefore, participants rated even the samples with much spectral degradation as similar to the reference. In the context of the literature, these findings are consistent with the findings of Smith et al (2013). In this study, it was found that participants who underwent testing on spectral resolution detection and discrimination preformed better with less electrical spread.

Another interesting interaction was seen between different genres and the other two factors. The quality ratings between different genres were found to be significantly different (Table 5). Folk Simple was found to be significantly different than all the other genres except the

Flute and Folk Complex. Folk Complex was found to be significantly different than all other genres except Folk Simple. Figure 14 breaks down the quality ratings of Rock Complex and Folk Complex. On the left graphs, quality ratings are shown by simulated electrical spread and spectral smearing. The vertical axis represents the quality ratings averaged between participants. The horizontal axis denotes the simulated electrical spread conditions. Each bar within an electrical spread condition signifies the smearing factor ranging between the reference and anchor (0.1). Again, there is the general trend for the quality scores to decrease with more spectral smearing, regardless of electrical spread or genre.

The graphs on the right denote the difference between the reference and anchor scores (horizontal axis) for each simulated electrical spread condition (vertical axis). When the difference graphs between genres are compared, a pattern emerges. The difference in ratings for Folk Complex are generally higher than the difference in ratings for Rock Complex. This suggests listeners had more difficulty hearing spectral differences in the Rock Complex music than in Folk Complex. The smaller range between the reference and anchor indicates participants had difficulty hearing spectral differences and rated the degraded samples as similar to the reference. One possible reason may be that the Rock Complex sample had electric guitar in it so there was already some distortion in the sample, and the additional processing (vocoding and cepstral analysis) exacerbated the distortion.

Figure 14 (left) Quality rating scores with spread condition on the horizontal axis and quality rating on the vertical axis. The smearing conditions are represented by the bars in each electrical spread condition.
(right) difference between the reference and 0.1 (anchor) ratings for each electrical spread condition. Electrical spr is on the vertical axis and the quality rating is on the horizontal axis.

Factor	df	F	P
Electrical Spread	3,39	19.7	P<0.001
Spectral Smearing	1.4, 18.1	255.7	P<0.001
Genre	2.4, 31.9	22.9	P<0.001
Spread * Genre	15, 195	3.6	P<0.001
Spread * Smearing	18, 234	13.8	P<0.001
Genre * Smearing	30, 390	25.7	P<0.001

Table 4: Main factors with degree of freedom, F-value, and P-value (a P-value of less than 0.05 is significant).

Table 5: Interaction of conditions within factors (a P-value of less than 0.05 is significant. * denotes significant difference).

Factor	Factor	P-Value
Spread		
	UNP*Neg40	0.02*
	UNP*Neg12	0.001*
	UNP*Neg6	P<0.001*
	Neg40*Neg6	0.025*
	Neg40*Neg12	1
	Neg12*Neg6	0.75
Smearing		
	Ref*0.7	0.014*
	Ref*0.5	0.001*
	Ref*0.4	P<0.001*
	Ref*0.3	P<0.001*
	Ref*0.2	P<0.001*
	Ref*0.1	P<0.001*
	0.7*0.5	0.2
	0.7*0.4	P<0.001*
	0.7*0.3	P<0.001*
	0.7*0.2	P<0.001*
	0.7*0.1	P<0.001*
	0.5*0.4	0.001*
	0.5*0.3	P<0.001*
	0.5*0.2	P<0.001*
	0.5*0.1	P<0.001*
	0.4*0.3	P<0.001*
	0.4*0.2	P<0.001*
	0.4*0.1	P<0.001*
	0.3*0.2	P<0.001*
	0.3*0.1	P<0.001*
	0.2*0.1	P<0.001*
Genre		
	Rock Simple* Rock Complex	1
	Rock Simple*Classical	1
	Rock Simple*flute	0.2
	Rock Simple*Folk Simple	0.001*
	Rock Simple*Folk Complex	P<0.001*
	Rock Complex* Classical	0.38
	Rock Complex* Flute	0.084
	Rock Complex* Folk Simple	P<0.001*
	Rock Complex* Folk Complex	P<0.001*
	Classical*Flute	1
	Classical* Folk Simple	P<0.001*
	Classical* Folk Complex	P<0.001*
	Flute*Folk Simple	0.1
	Flute*Folk Complex	0.005*
	Folk Simple* Folk Complex	1

General Discussion and Conclusion

Overall, the results of the study show that with less simulated electrical spread music quality perception significantly improved, even when environmental factors like spectral resolution and acoustic properties were varied. These results support the idea that the focused stimulation processing strategy may improve music perception compared to monopolar stimulation. While these findings are interesting, there is more research that must be done due to limitations in the experiment.

This study used simulations of CIs with normal hearing individuals which allowed us to carefully control specific differences in processing (c.f., Pals et al, 2013; Arehart et al, 2014; Tamati & Baskent, 2018). This control is both a strength and a limitation of the experiment. The limitation is that the simulation may not include all the factors that affect CI listening. One next step will be to compare the results from this experiment to the results that Cochlear Ltd is collecting from CI users. This comparison may yield differences in results between the two groups due to differences in the auditory system or because of the effects of sensory adaptation in CI users from listening to sound processed through their device all the time.

Another factor that should be further researched is genre. A systematic analysis of acoustic differences across genres might provide insights into the genre differences observed in this study. In addition, the genres tested in this study spanned a small range of all the music people listen to. Further research could be done with more genres like heavy metal, rap, or acapella to give a wider representation to the acoustic differences in music.

The results from this study have implications for music perception for CI users. The continued research of CIs and music is vital because music is integral to the human experience (Dasa, 2018; North & Hargreaves, 2003; Papinczak et al, 2015). Although there is still room for improvement for music perception in CIs, CI users still have positive experiences with music. This is evident after hearing stories from Rachel Kolb, a young woman who received a CI when she was 21years old (Kolb, 2018). After receiving her implant, Rachel began to explore the world of music. At her first symphony concert, Rachel described being jolted by the drums and feeling the violin's melody pierce her chest. She broke her belief that "deaf dancing" was an oxymoron by going dancing at a club with her friends. Through her experiences with music, Rachel came to understand that the "celebration of feeling, motion, sensation, and language was what mattered when [she] experienced music". Rachel may not hear the same details in music that her friends with normal hearing do, but she experiences the magic of music in a way that is significant to her.

Acknowledgements

I would like to thank my committee for their support throughout the thesis process. I would especially like to thank Dr. Arehart and Dr. Croghan for guiding me through the research process and providing moral support. Thank you to Cochlear Ltd and University of Colorado Undergraduate Research Opportunity Program (UROP) for providing funding for this experiment. Finally, I would like to thank my family and friends for their encouragement throughout my college journey.

Reference

Arehart, K. H., Croghan, N. B. H., & Muralimanohar, R. K. (2014). Effects of age on melody and timbre perception in simulations of electro-acoustic and cochlear-implant hearing. *Ear and Hearing*, *35*(2), 195–202. https://doi.org/10.1097/AUD.0b013e3182a69a5c

Aronoff, J. M., & Landsberger, D. M. (2013). The development of a modified spectral ripple test. *The Journal of the Acoustical Society of America*, *134*(2), EL217-EL222. https://doi.org/10.1121/1.4813802

Bieber, J. (2009). Baby [recorded by Justin Bieber] on *My World 2.0* [CD]. Atlanta, United States: Triangle Sound Studios.

Brant, J. A., Eliades, S. J., Kaufman, H., Chen, J., & Ruckenstein, M. J. (2018). AzBio Speech Understanding Performance in Quiet and Noise in High Performing Cochlear Implant Users. *Otology & Neurotology*, *39*(5), 571. https://doi.org/10.1097/MAO.0000000000001765

Croghan, N. B. H. (2013). *Perceived Quality of Recorded Music Processed through Compression Hearing Aids* (Ph.D.). University of Colorado at Boulder, United States -- Colorado. Retrieved from https://search.proquest.com/docview/1436138535/abstract/C063465B41C346FAPQ/1

Croghan, N. B. H., Duran, S. I., & Smith, Z. M. (2017). Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility. *The Journal of the Acoustical Society of America*, *142*(6), EL537-EL543. https://doi.org/10.1121/1.5016044

Dassa, A. (2018). Musical Auto-Biography Interview (MABI) as promoting self-identity and well-being in the elderly through music and reminiscence. *Nordic Journal of Music Therapy*, *0*(0), 1–12. https://doi.org/10.1080/08098131.2018.1490921

Diamond, N. (1969) [recorded by Neil Diamond]. On *Dig In* [record] Memphis, United States: Uni/MCA

Dritsakis, G., Besouw, R. M. van, & Meara, A. O. (2017). Impact of music on the quality of life of cochlear implant users: a focus group study. *Cochlear Implants International*, *18*(4), 207–215. https://doi.org/10.1080/14670100.2017.1303892

Eggermont, J. J. (2017). Acquired hearing loss and brain plasticity. *Hearing Research*, *343*, 176–190. https://doi.org/10.1016/j.heares.2016.05.008

Fishman, K. E., Shannon, R. V., & Slattery, W. H. (1997). Speech Recognition as a Function of the Number of Electrodes Used in the SPEAK Cochlear Implant Speech Processor. *Journal of Speech, Language, and Hearing Research*, *40*(5), 1201–1215. https://doi.org/10.1044/jslhr.4005.1201

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America*, *110*(2), 1150–1163. https://doi.org/10.1121/1.1381538

Garnham, C., O’Driscoll, M., Ramsden, R., & Saeed, S. (2002). Speech Understanding in Noise with a Med-El COMBI 40+ Cochlear Implant Using Reduced Channel Sets. *Ear and Hearing*, *23*(6), 540.

Gifford, R. H., Shallop, J. K., & Peterson, A. M. (2008). Speech Recognition Materials and Ceiling Effects: Considerations for Cochlear Implant Programs. *Audiology and Neurotology*, *13*(3), 193–205. https://doi.org/10.1159/000113510

Gold T. (1948) Hearing. II. The physical basis of the action of the cochlea. Proc R Soc Lond B Biol Sci**135**(881):492–498.

Henry, B. A., Turner, C. W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. *The Journal of the Acoustical Society of America*, *118*(2), 1111–1121.

Kolb, R. (2018, January 20). Opinion | Sensations of Sound: On Deafness and Music. *The New York Times*. Retrieved from https://www.nytimes.com/2017/11/03/opinion/cochlear-implant-sound-music.html

Landsburger, D. (July, 2017). *Music Enjoyment in Single-Sided Deafened Patients: The Synergistic Effect of Electric and Acoustic Stimulation*. Poster session presented at CIPediatrics San Francisco, San Francisco, CA.

Limb, C. J., & Roy, A. T. (2014). Technological, biological, and acoustical constraints to music constraints to music perception in cochlear implant users. *Hearing Research*, *308*, 13–26. https://doi.org/10.1016/j.heares.2013.04.009.

Litvak, L. M., Spahr, A. J., Saoji, A. A., & Fridman, G. Y. (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *The Journal of the Acoustical Society of America*, *122*(2), 982–991. https://doi.org/10.1121/1.2749413

Litvak, L. M., Spahr, A. J., Saoji, A. A., & Fridman, G. Y. (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *The Journal of the Acoustical Society of America*, *122*(2), 982–991. https://doi.org/10.1121/1.2749413

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences*, *103*(49), 18866–18869. https://doi.org/10.1073/pnas.0607364103

NIDCD. (2015, August 18). Retrieved September 19, 2018, from https://www.nidcd.nih.gov/health/how-do-we-hear

NIDCD. Cochlear Implants. (2015, August 18). Retrieved September 24, 2018, from https://www.nidcd.nih.gov/health/cochlear-implants

Oghalai, J. S. (2004). The cochlear amplifier: augmentation of the traveling wave within the inner ear. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *12*(5), 431–438. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1315292/

Oxenham, A. J. (2008). Pitch Perception and Auditory Stream Segregation: Implications for Hearing Loss and Cochlear Implants. *Trends in Amplification*, *12*(4), 316–331. https://doi.org/10.1177/1084713808325881

Pals, C., Sarampalis, A., & Baskent, D. (2013). Listening Effort With Cochlear Implant Simulations. *Journal of Speech, Language and Hearing Research (Online); Rockville*, *56*(4), 1075–1084. https://doi.org/http://dx.doi.org.colorado.idm.oclc.org/10.1044/1092-4388(2012/12-0074)

Papinczak, Z. E., Dingle, G. A., Stoyanov, S. R., Hides, L., & Zelenko, O. (2015). Young perception in cochlear implant users. *Hearing Research*, *308*, 13–26. https://doi.org/10.1016/j.heares.2013.04.009

Plack, C. J. (2013). *The Sense of Hearing: Second Edition* (2 edition). New York: Psychology Press.

Roy, A. T., Jiradejvong, P., Carver, C., & Limb, C. J. (2012). Assessment of sound quality perception in cochlear implant users during music listening. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *33*(3), 319–327. https://doi.org/10.1097/MAO.0b013e31824296a9

Roy, A. T., Jiradejvong, P., Carver, C., & Limb, C. J. (2012). Musical Sound Quality Impairments in Cochlear Implant (CI) Users as a Function of Limited High-Frequency Perception. *Trends in Amplification*, *16*(4), 191–200. https://doi.org/10.1177/1084713812465493

Roy, A. T., Penninger, R. T., Pearl, M. S., Wuerfel, W., Jiradejvong, P., Carver, C., … Limb, C. J. (2016). Deeper Cochlear Implant Electrode Insertion Angle Improves Detection of Musical Sound Quality Deterioration Related to Bass Frequency Removal. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *37*(2), 146–151. https://doi.org/10.1097/MAO.0000000000000932

Smith, Z. M., Parkinson, W. S., & Long, C. J. (2013). Multipolar current focusing increases spectral resolution in cochlear implants. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2796–2799). https://doi.org/10.1109/EMBC.2013.6610121

Tamati, T. N., Janse, E., & Başkent, D. (2018). Perceptual Discrimination of Speaking Style Under Cochlear Implant Simulation. *Ear and Hearing*. https://doi.org/10.1097/AUD.0000000000000591

Won, J. H., Drennan, W. R., & Rubinstein, J. T. (2007). Spectral-Ripple Resolution Correlates with Speech Reception in Noise in Cochlear Implant Users. *JARO: Journal of the Association for Research in Otolaryngology*, *8*(3), 384–392. https://doi.org/10.1007/s10162-007-0085-8

Zhao, K., Bai, Z. G., Bo, A., & Chi, I. (2016). A systematic review and meta-analysis of music therapy for the older adults with depression. *International Journal of Geriatric Psychiatry*, *31*(11), 1188–1198. https://doi.org/10.1002/gps.4494