A real-time cochelar implant simulator - design and evaluation

Steinhauer, Christina; Sønderbo, Tobias; Paisa, Razvan

Published in:
SMC 2023 - Proceedings of the Sound and Music Computing Conference 2023

DOI (link to publication from Publisher):
10.5281/zenodo.10063009

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Steinhauer, C., Sønderbo, T., & Paisa, R. (2023). A real-time cochelar implant simulator - design and evaluation. In R. Bresin, & K. Falkenberg (Eds.), *SMC 2023 - Proceedings of the Sound and Music Computing Conference 2023* (pp. 410-418). Sound and Music Computing Network. https://doi.org/10.5281/zenodo.10063009

A REAL-TIME COCHELAR IMPLANT SIMULATOR - DESIGN AND EVALUATION

Christina Steinhauer
Aalborg University Copenhagen
cstein22@student.aau.dk

Tobias Sønderbo
Aalborg University Copenhagen
tlso22@student.aau.dk

Razvan Paisa
Aalborg University Copenhagen
rpa@create.aau.dk

ABSTRACT

This article describes the implementation of a flexible real-time Cochlear Implant (CI) simulator, and it's preliminary evaluation set to investigate if a specific set of parameters can simulate the musical experience through CIs using Normal Hearing (NH) subjects. A Melodic Contour Identification (MCI) test is performed with 19 NH subjects to identify melodic contours processed by the simulator. The results showed that the participants had a decrease in precision in determining musical contours as the intervals between notes decreased, showing that the reduced spectral resolution increases the difficulty to identify smaller changes in pitch. These results fall in line with other studies that perform MCI tests on subjects with CI, suggesting that the real-time simulator can mimic the reduced spectral resolution of a CI successfully. This study validates that the implemented simulator, using a pulse-spreading harmonic complex as a carrier for a vocoder, can partially resemble the musical experience had by people with hearing loss using CI hearing technology. This suggests that the simulator might be used to further examine the characteristics that could enhance the music listening experience for people using CIs.

1. INTRODUCTION

A cochlear implant (CI) is a device that restores part of the auditory sensations for those who suffer severe-to-profound hearing loss. The implant stimulates the auditory nerve directly in the cochlea, based on auditory input captured by the external sound processor [1], and it can provide good speech comprehension in quiet environments [2]. However, CI listeners experience poor music perception, both in terms of self-reported music enjoyment and objective perceptual abilities, scoring significantly lower than NH subjects [3–5]. This stems from a deficit of pitch and timbre perception, which is limited by the spectral resolution of the CI [5, 6], as well as an insufficient providence of input dynamic range to cover the wide amplitude range of music [7]. Proper pitch perception is crucial to understanding the harmonic complex tones produced by musical instruments or when distinguishing sources, for example in situations with multiple talkers [6]. Since music is an important part of social gatherings and mood regulation, the limitations of music perception in CI individuals may affect their well-being and quality of life [8], demanding that further research is needed to improve upon this.

The end results of auditory performance of the CI receiver can be affected by multiple factors such as the choice of device, surgical quality, duration of hearing loss (HL), whether HL occurred pre- or postlingual, the quality of recovery from surgery, rehabilitation practices, and even more [9]. This results in a large variation in the individual experience that applies to the perception and appreciation of music. For instance, studies suggests that early deafened, late implant users appreciates music more than postlingual users, although there seemed to be no significant difference between the two groups in identifying musical contours [10].

This study proposes an implementation of a CI simulator application that processes sound in real-time, and can be adjusted with user-configurable parameters with immediate effect. This flexibility encourages further exploration of various auditory perceptions experienced by CI users and discussion about how the combination of parameters affects the experience. The simulator incorporates three carriers: sine, noise, and pulse-spreading harmonic context (PSHC), the latter also being evaluated in a melodic contour identification (MCI) test. Here, we assessed nineteen NH subjects' ability to perceive musical contours by the simulator, with similar configured parameters to those outlined in Karoui et. al. [11]. In doing so, we wish to draw parallels to an MCI test conducted with CI subjects Galvin et al. [12] and to discuss if the parametric setup can infer anything about the musical perception of CI users.

2. BACKGROUND

To get a better understanding of the different parameters and shortcomings of CI technology, there is an increasing interest in simulating the auditory perception of CI listeners. Channel vocoders are typically used to acoustically simulate CIs, usually with a number of sinusoidal or noise-band carriers that simulate the electrical pulse trains from the implanted electrodes [11]. The vocoder approximates the signal chain found in cochlear implants, as electrodes can be thought as limited bandwidth carriers, coupled directly to the auditory nerve.

A vocoder (voice coder-decoder) analyzes a signal by separating it into a number of channels and then extracts the low frequency envelope of each channel. The vocoder recreates the original signal using the extracted envelopes to modulate either noise or carrier waves for each separate channel and then summarizes these into a final, re-synthesized signal.

Vocoder based simulators are not a novelty, having been used in several studies, these have been mostly evaluated in terms of speech intelligibility spatial localization [13–15]. In 2006 Poissant et al. [15] started to look at the effects of reverberation on speech in quiet scenario with multiple vocoder configurations, concluding that systems with a lower number of channels exhibit an exaggerated vulnerability to reverberation. Similarly, Whitmal et al. [16] compares two vocoder channels in terms of sentence identification, concluding that noise based vocoders perform worse than tone based ones, mainly due to their intrinsic temporal modulation that can interfere with the temporal fluctuations carrying speech cues. More recently, Jain and Ghosh [17] set to investigate the effect of several simulator parameters on speech quality and intelligibility using both subjective and objective tests.

Another approach to investigate the validity of CI simulators proposed by several studies [18, 19] is to compare the electric and acoustic stimulation in CI subjects with single-sided deafness (SSD). Both these studies found substantial variability in selection of the most similar sounding simulation to the CI ear. A recently introduced pulsatile carrier, pulse spreading harmonic complex (PSHC [20]), aims to mitigate the limitations of sinusoidal and noise carriers for CI simulations outlined by [16]; specifically, that sinusoids cannot simulate the broad spread of excitation produced by a CI electrode, [21], and that noise contains intrinsic modulations that are absent in CIs. PSHC is broadband and intrinsic modulations can be minimized by its pulse rate [20]. When evaluated by SSD-CI subjects using tone, noise and PSHC-based vocoders, it was found that the PSHC ones were judged more similar to the CI, than the other two cases [11, 22].

One thing that all aforementioned simulators have in common is that they do not allow for real-time control of parameters, or some work only with pre-recorded sounds, severely limiting the possibility of exploration and customization. Cavdir et al. [23] mentioned that when working with accessible technology it is important to acknowledge the individual characteristics and engage in an open discussion with the target group, thus, in the case of CI simulators it is important to be able to adjust both input signals and the technical characteristics of the system.

3. MATERIALS AND METHODS

In this study, we developed a CI simulation application which aims to simulate the auditory experience of CI users in real time. It may be configured based on several parameters, allowing us to consider several different perceptions of CI. These parameters include amount of channels, carrier type, frequency range of the simulated implant, individual channel gain and overall compression (threshold and makeup gain). The main layout of the application and parameters are displayed in Figure 1.

3.1 Software

The CI Simulation application was based on *JUCE* [1] as the framework due to its run-time performance benefits, which were required for real-time processing. - and an interactive UI allowing configuration of several parameters in real-time. The MCI test was implemented in Windows Forms (*WinForms .NET*). The MCI test results was analysed in MATLAB and visualized with PowerBI.

3.2 Cochlear Implant Simulation Processor

The processor is based on a vocoder to represent the CI, with each channel corresponding to a subset of the electrodes inserted into the cochlear. The base design of the vocoder is implemented as described by Karoui et al. [11]. The implementation of the CI simulation processor is designed to be configurable, which can be directly interacted using the UI components of the application. As the application executes the simulation in real time, the processor fetches the parameter values from a state that contains the most recent settings based on the configurable components. When the processor receives an *audio block*, it is duplicated into N amount of channels, where the number (N) is user configurable. The channels are then processed through three stages: *preprocessing*, *analysis* and *reconstruction* as shown in Figure 5. After being processed through these three stages, the output signal is then compressed with a configurable threshold. Furthermore, there exist additional options that allows controlling the gain of each channel and overall output signal.

3.2.1 Preprocessing

In the *preprocessing* stage, the signal is filtered with a band-pass filter to limit the frequency range. The upper and lower frequencies are configurable, but have a default value of 250Hz-4500Hz to match the ones in Karoui et. al. [11] and ensure comparability between studies. In the other study, the limited frequency range served the purpose of being able to control the input stimuli passed to SSD-CI subjects, by low-pass filtering all CI stimuli so that the stimuli delivered to the two ears always had the same bandwidth. Furthermore, the upper limit is just over the highest fundamental frequency available on a piano, thus affording a broad selection of musical notes to be reproduced. An example of an input signal used in the experiment described in 3.3 can be seen in Figure 2; it is a simple piano melody that consists of five tones, starting in A4 and increasing in pitch by two semitones each time. The pre-filtered signal in the time domain (A) has approximately even amplitude peaks for all tones, in the frequency domain (B) the fundamental frequency (f_0) of the five tones span approximately between 400Hz and 750Hz. Post-filtered signal has been filtered with a sixth-order Butterworth band-pass described in Figure 3, which in this case represents a single channel in the vocoder; this will be elaborated upon in 3.2.2. As a result, on time (C) and frequency domain (D) the amplitude peaks decrease for each tone, as the frequencies pass beyond the upper cutoff frequency of the filter (500Hz).

[1] https://juce.com/

Figure 1. Screenshot of the application in action with a microphone as input source device. The upper section of the layout consists of two displays, visualizing the frequency spectrum of the raw input and processed output; with each channel displayed individually in the output spectrum. The lower section of the layout constitute the user-configurable parameters, and can be modified to adjust the output in real-time. Use **(A)** to select number of channels, **(B)** to enable and disable carrier types, and control the gain of each carrier, **(C)** to control the gain of each individual channel, and **(D)** to adjust the upper and lower cut-offs of the frequency range in which the band-pass filters will be distributed between.

The envelope (E) is extracted with half-wave rectification followed by a second order Butterworth low-pass filter.

3.2.2 Analysis

In the *Analysis* stage, the temporal information for each channel is processed and extracted. Sixth-order Butterworth band-pass is used to filter each channel which greatly reduce spectral information, but preserves the temporal envelope cues in each band [24].

An example of a sixth-order Butterworth is shown in Figure 3 with cutoff frequencies at 250Hz and 500Hz. The effect of the filter used on a signal of a simple piano melody can be seen in Figure 2.

The lower and upper frequencies of the Butterworth filters used in the *analysis* stage are calculated using the Greenwood function [25] within the frequency range used in the *preprocessing* stage (3.2.1). The *N** channels covers their own section of the frequency spectrum with the theoretical width of excitation along the basilar membrane [26] (see Figure 4). The envelope is extracted with half-wave rectification, followed by a second order Butterworth low-pass filter. Similar to the approach taken by Mesnildrey et al. [21], the cutoff frequency for the low-pass filter is equivalent to the pulse rate of the PSHC carrier divided by two. This will be further elaborated in the Reconstruction stage (3.2.3). The cutoff frequency for the low-pass filter has a upper limit of $200Hz$ for cases wherein half the frequency of the pulse rate is greater than the aforementioned value.

3.2.3 Reconstruction

In the *Reconstruction* stage, each channel is synthesized to produce acoustic simulation. The envelope of each channel is modulated using one or more of three carriers: sinusoidal (SINE), Gaussian noise (NOISE) or PSHCs. The carrier modulation resembles the procedure used by Karoui et al. [11]. As such, NOISE carrier generates noise based on Gaussian distribution. SINE carrier generates a sinusoid with the center frequency of the corresponding bandpass. The PSHC carrier is implemented as described in [21]. Therefore, all PSHC carriers have a fundamental frequency of 0.3 Hz and the pulse rate is frequency dependant for each channel, in order to limit intrinsic modulations. The optimal pulse rate is calculated using equation 1, which is a second order polynomial fit between center frequency and pulse rate derived from *Table 1* in [21]:

$$y = 37 + 151x + 0.17x^2 \tag{1}$$

where x is the center frequency and y is the optimal pulse rate.

Finally, all PSHCs were filtered with a gammatone filter as implemented in [27]), which were initialized with the same center frequencies of the band of the analysis. This optimizes the carrier such that their intrinsic modulations after auditory filtering, with the optimal pulse rate, showed

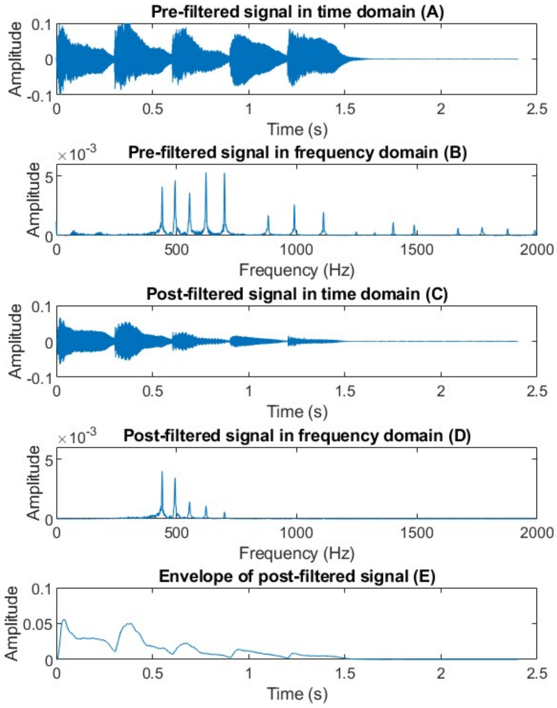

Figure 2. Input signal in different pre-processing stages in both time and frequency domains

smaller internal crest factors than other carriers with equivalent bandwidths [20]. After amplitude modulation, the output signal is filtered once again using the corresponding Butterworth band-pass filters from the Analysis stage. The implementation of the simulator does not include configurable ERB (equivalent rectangular bandwidth) mismatch.The authors in [11] found that participants overall had a preference for no ERB mismatch, as such this configuration was omitted from this simulator. For the experiment described in 3.3, only PSHC carries have been used, as this configuration is documented to be closer to CI's as highlighted in [11, 22].

3.3 Melodic Contour Identification Evaluation

The MCI test was based on the procedure described by Galvin Et al. [12, 28, 29], with the simulator parameters set to match those of the experiment described by [11]. The goal was to investigate whether our simulator can perform similarly to the one validated by bi-modal users, as presented in [11], when listening to music.

3.3.1 Subjects

There were 19 participants in the MCI test. All participants volunteered for the test, and no information regarding participants was recorded. Furthermore, none of the participants had any hearing loss based on self report.

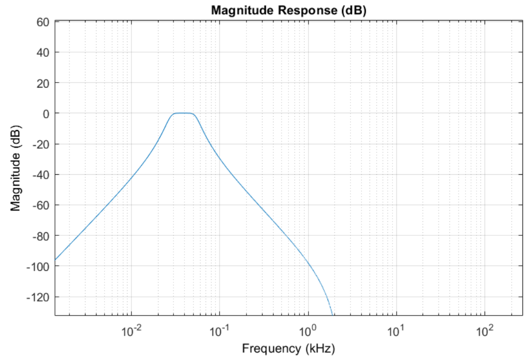

Figure 3. Magnitude response of a sixth-order Butterworth bandpass filter with 250 Hz and 500 Hz as lower and upper cutoff frequencies respectively. The Frequency (kHz) axis is logarithmic.

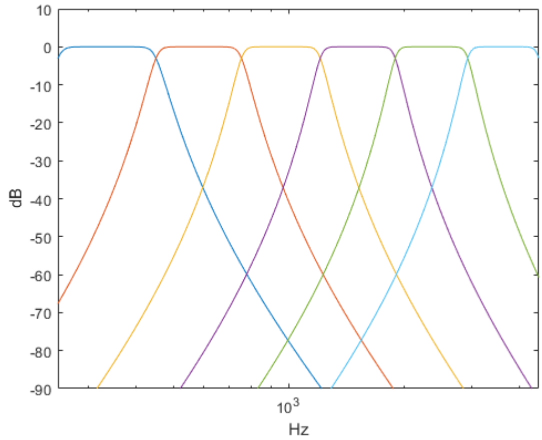

Figure 4. Analysis Butterworth bandpass filters for 6 channels spanning from 250-4500Hz, distributed with the greenwood function. The x-axis (Hz) is logarithmic.

3.3.2 Stimuli

Stimuli consisted of equal amplitude virtual grand piano (Ableton Live 11 Grand Piano collection [2]) triggered by MIDI notes which has been processed using the CI simulator. Piano sounds were used as they produce overtones giving the notes harmonic components. This makes some notes audible even if their fundamental frequency is below cutoff frequency of the band-pass filter in the simulator. The simulator was configured with the parameters described in [11], due to the PSHC vocoder being judged most similar compared to SINE and NOISE vocoders with these configurations. It was processed through six channels spanning from 250Hz-4500Hz and synthesized with PSHC as carrier type. In order to obtain results as comparable as possible with the ones in [11], no compression was added to the stimuli.

3.3.3 Procedure

Similar to previous studies investigating contour identification for CI users [12, 28, 29], the experiment described here

[2] www.ableton.com

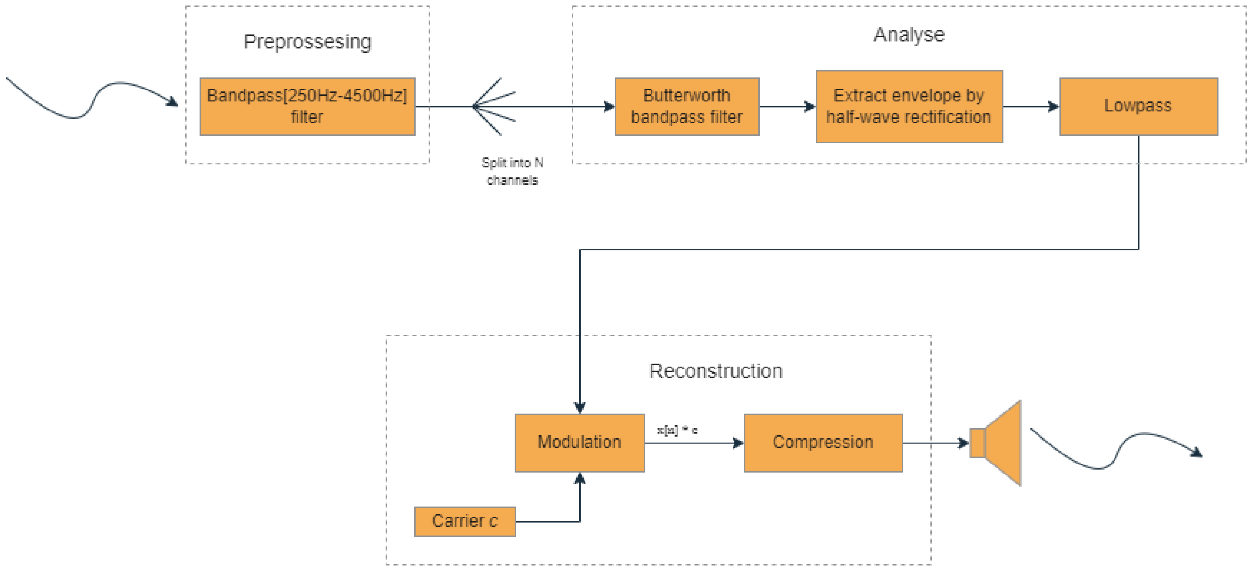

Figure 5. Block diagram of the three stages of the CI simulation processor

used five notes of equal duration following musical intervals. Each note was played for the duration of 250 ms with 50 ms of rest between notes. Each set of notes used either two, three or five semitones separation between each other, which differs from the procedure by Galvin et al. [12] as their evaluation used all combinations of one to five semitone intervals. This was done to simplify the experiment, as an initial evaluation of the CI simulator and PSHC carrier type. The melody played for a set is either *Rising, Falling* or *Flat* as depicted in Figure 6. The root note was the starting note for the contours in three different octaves (A3, A4 and A5). This leads to a total of twenty-one different combinations of melodies. Each combination was repeated three times in a random order for each participant, resulting in a total sixty-three test cases. The participants were not informed about the number of melodies and contours, or the number of repetitions for each case, and they were only allowed to hear each contour only once.

A custom PC application was used to perform the test. The application played each contour based on an input from the participant. The participant sat in front of a laptop computer in an empty classroom, and could select the interpreted contour from the three options presented (Falling, Flat, Rising) as seen in Figure 6. It took each participant five to ten minutes to complete the test. The MCI test results of all participants were recorded and stored in a database with relevant metadata (contour type, semitone interval, octave and selected contour). The sounds were played to the participant at a comfortable level using a set of Bose QC700 noise-cancelling headphones, with noise cancellation set to the highest level.

In summary, 21 different combinations of contours were tested: *contour*{Falling, Rising} * *octave*{A3, A4, A5} * *semitone*{2, 3, 5} and *contour*{Flat} * *octave*{A3, A4, A5}.

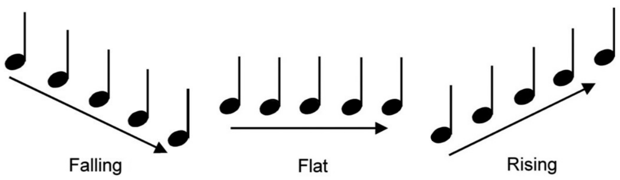

Figure 6. Depiction of Falling, Flat, and Rising melodic contours.

4. RESULTS

This section will cover the results of the MCI test. The test included 19 participants for a total of 1197 answers. Of the answers submitted 511 of them were incorrect and 686 of them were correct resulting in an average performance of 57.3% accuracy in determining contours with a standard deviation of 12.8% [3]. A detailed overview of the amount of correct answers per contour in percentage is detailed in Table 1.

A Wilcoxon signed-rank test was conducted to determine the effect of accuracy based on observations of pairs in their respective category (Octaves and Semitones). The nineteen participants' answers were summarized based on each category and paired up to test if there were any significant difference between the pairs. The significance level

[3] A bug was discovered in the program used to test the participants. This caused one contour to be played four times instead of three - and a random other contour to be played two times instead of three. Because of this, a set of seventeen participants had an extra flat contour in their test, replacing a falling contour for a subset of eleven participants, and replacing a rising contour the remaining subset of six participants. Every participant has been exposed to each combination of contour, octave and semitone interval, and has evaluated a total of sixty-three contours.

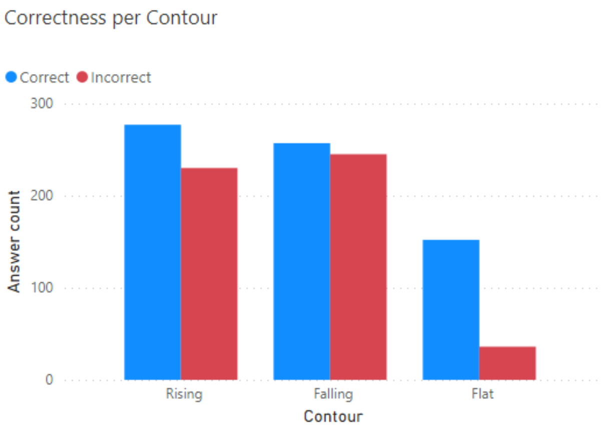

Figure 7. Results of the MCI test shown as correct and incorrect answers by contour. The most significant difference in amount of correct answers is the Flat contour with 152 correctly identified and 36 incorrect.

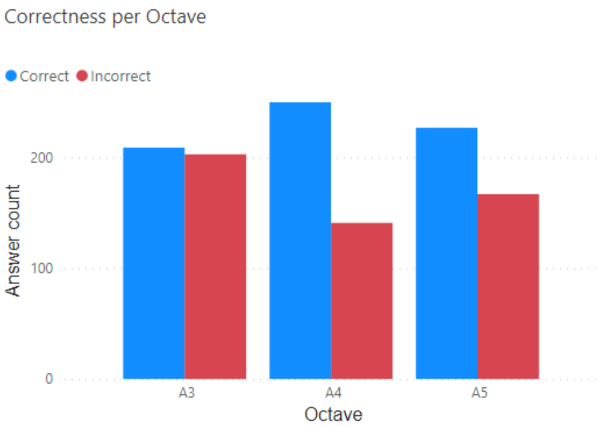

Figure 8. Results of the MCI test shown as correct and incorrect answers by octave. The plot shows that the A4 octave has the largest proportion of correctly identified answers and A3 has the lowest.

was set to $\alpha = 0.05$ and compensated using Bonferroni correction $\alpha/6 = 0.008$. The pairs and results are shown in Table 2.

All participants performed equal or above chance level ($\geq$ 33%) with the worst two participants settled just at chance level at 33% accuracy and the best one was 74.6% correct. Distribution of correct and incorrect answers in their respective category can be seen in Figure 7 for the contour types, Figure 8 for octaves and Figure 9 for semitone intervals.

5. DISCUSSION

5.1 Analysis of Results

The distribution of correct answers shown in Figure 11 shows that participants are able to achieve an accuracy that is equal to or higher than chance level. Plotting the distribution of correct and incorrect answers by the different

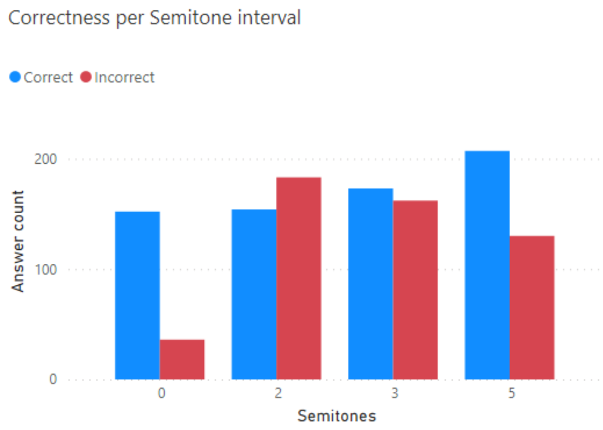

Figure 9. Results of the MCI test shown as correct and incorrect answers by Semitone interval between each note in the contours. The interval of 0 corresponds to all of the flat contours. The 0 semitone difference shows the largest difference with most of answers being correct. From 2 to 5 semitones the proportion of correct answers increase with only the 2 semitone difference showing less correctly identified contours than correctly identified.

	%	A3	A4	A5	semi-2	semi-3	semi-5
Falling	51.20	34.13	62.87	56.55	42.86	52.38	58.43
Flat	80.85	75.00	76.36	92.98	-	-	-
Rising	54.63	56.21	60.94	46.75	48.52	50.90	64.33
Mean	57.31	50.00	65.79	56.89	45.03	53.56	60.52
SD	12.80	13.72	23.29	18.81	13.47	18.49	20.11

Table 1. Correct answers in percentage of the contour combinations.

types of contours shows some of the key areas where participants are more consistent in correctly identifying the contours. Figure 7 shows that participants are most accurate with flat contours with an accuracy of 80.8%. The flat contour was also the most commonly given answer for a contour as shown in Figure 10. This implies that participants struggled to discerning the difference in pitch between notes. As shown in Figure 9, participants performed best with zero semitone interval (flat contours), and for the remaining semitone intervals, there is an improvement in accuracy with each increase in semitone interval. The interval of two semitones is the only interval which has less correctly identified contours than incorrect. There is a statistically significant median increase in accurate answers between two (median of 9) and five (median of 11) semitones, $p = .002$. This could be explained by the difference in pitch being the smallest for the three interval types causing participants to interpret Rising and Falling as Flat. The percentage of answers given as "*flat contour*" decreases as the semitone interval increases, as shown in Table 3, implying that it becomes easier to discern the difference in pitch as the interval between notes increase.

Of the three different octaves, A3 had the lowest accuracy. The frequency of the root note A3 is 220Hz which puts it outside of the cutoff frequency for the initial band

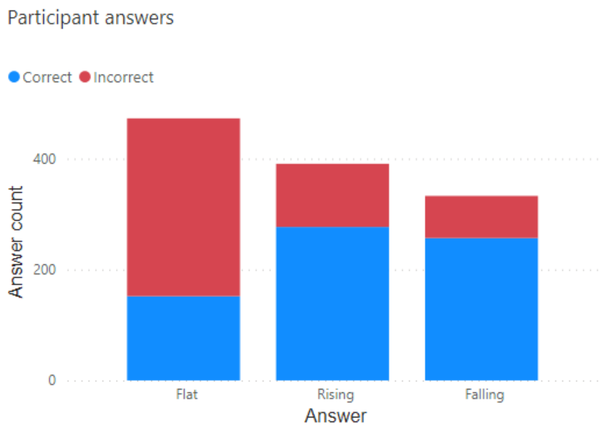

Figure 10. Distribution of answers by the participants for all contours. The most common selected contour was Flat. Falling was the least selected contour.

X	Y	p	p ¡ 0.008
Octave A3	Octave A4	0.07	No
Octave A3	Octave A5	0.5	No
Octave A4	Octave A5	0.3	No
Semitone 2	Semitone 3	0.2	No
Semitone 2	Semitone 5	0.002	Yes
Semitone 3	Semitone 5	0.02	No

Table 2. Wilcoxon signed rank test for zero median with paired samples within their own category of contour test cases. The test pairs that has a statistically significant median difference is the contour pairs Semitone 2 and 5, Semitone 3 and 5, Flat and Rising, and Flat and Falling.

pass filter used. This causes all of the notes in a falling and flat contour using A3 as root to be filtered by the band pass filter, with only the rising contour having notes that are not filtered. By filtering the results by the A3 octave, only the Rising and Flat contours show a positive difference in correctly identified contours. For the falling contour in A3, 110 out of 167 were not correctly identified, this makes up 45% of all incorrect answers for falling contours. Given that the notes for falling and flat both fall outside the lower limit of the frequency range (250Hz), this may increase the difficulty discerning the the difference between the two contours. This also explains why rising contours has a higher percentage of correct answers compared to falling in A3.

The Octave with the highest count of correct answers is A4, with a positive difference in the amount of correct answers compared to incorrect across all contours. The frequency of the root A4 is 440Hz, putting it 10Hz below the upper cutoff frequency of the first channel band, making rising contours span up to 3 channels depending on semitone interval. For falling contours, semitone intervals of three and five will go below the limit of the frequency range. Results for falling and rising contours are similar, which could suggest that the initial notes of the contours are the most important for determining the type of contour.

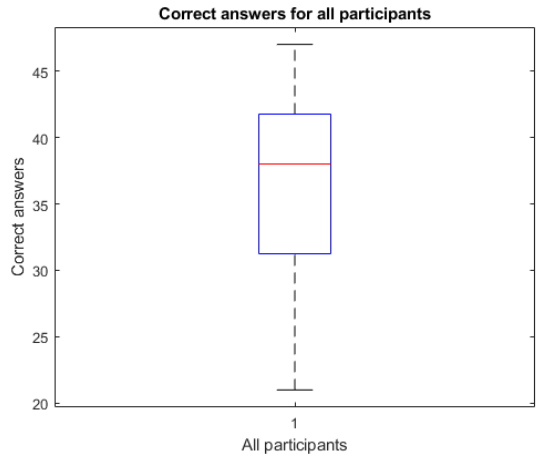

Figure 11. Box plot of the amount of correct answers per participant. The median of thirty-eight (38) is marked by the red line and the 25th percentile and the 75th percentile marked by horizontal blue lines showing 31.25 and 41.75 respectively. The lowest scoring participant had a score of twenty-one (21) and highest scoring participant had a score of forty-seven (47).

%	Semitone 2	Semitone 3	Semitone 5
Flat	40.94	30.01	24.33

Table 3. Percentage of answers given as Flat per semitone interval

Looking at the octaves themselves, the p-values shown in Table 2 does not suggest that the octaves on their own influence the amount of correct answers. The only pair observed that produces a p-value that approaches 0.05 is A3 and A4 ($p = .07$), but as discussed earlier, this is likely due to other factors.

5.2 Simulated vs. Actual

The results of the MCI test show a mean result similar to the one performed in the reference paper [12], in which nine CI subjects performed a MCI test. Nevertheless, they show a much wider standard deviation in results due to participants using their clinically assigned CI, which includes several models as well as users having different backgrounds with their CI's. The detailed results show a similar issue in determining contours with small intervals between notes, with the accuracy increasing with increases in semitone intervals. The results with octaves in [12] vary from user to user with some users performing better in some octaves that others. The results of the test performed on this simulator favors the A4 octave, but configured differently the results might be more in line with that of actual CI models. The CI users tested in [12] showed a similar tendency as the participants of this test, with the most frequent answer being *flat (contour)* and the least one being *falling (contour)*. Comparing the configuration of this simulator to the characteristics of the CI models used in [12], the one described in this article uses a shorter frequency range than that of the CI models. The lower frequency limit go as low as 120Hz and the upper frequency limit as high as 10853Hz. Configuring the simulator to have a similar range could improve performance as it would solve some of the previously mentioned issues with notes falling outside the band-pass filter. Results of both tests support that many contours are misinterpreted as flat contours. This is a result of the short intervals between notes, that when processed, becomes difficult to discern due to the limited spectral resolution.

5.3 Research Limitations

The limited number of participants for the test reduces the amount of confidence in some of the results. The results produced from the test show a significant deviation in results between participants in certain categories. The limited amount of participants also makes determining outliers difficult. The test would benefit from a larger set of participants, making results more reliable.

5.4 Future Research Suggestions

Galvin et al. [12] showed that accuracy in determining contours of its participants increased with training. Rerunning the test with the simulator with a fixed set of NH subjects could investigate if this is also the case for a simulated environment. Rerunning the test with different sets of parameters could provide insight into which parameters impact the ability to perceive musical contours. Furthermore, in [11] it is implied that the PSHC carrier was judged more similar than a SINE or NOISE carrier for vocoder-based simulation of CI by SSD-CI subjects. It could be interesting to evaluate if mixing different set of carriers was judged more or less similar to their CI ear.

Lastly, this study did not utilize the real-time feature of the CI simulator. Many variables must be taken into consideration when conducting an experiment in real time in terms of stimuli requiring a separate study investigating the interaction between different parameter configurations. Nevertheless, allowing participants to select their desired configurations could open up for interesting and individual results. Lastly, the experiment described in this article aimed to be as comparable as possible to the one presented by [11], in order to understand how our simulator performs relative to the reference one, therefore the real-time feature was not necessary for the MCI evaluation.

6. CONCLUSION

The results of the MCI test show that NH listeners show similar tendencies to CI listeners when identifying melodic contours that has been simulated. In general it becomes more difficult to discern differences in pitch in melodic contours, with test participants often false classifying contours as flat. This becomes increasingly clear as the interval between notes decrease. The test performed could benefit from having more participants, as some of the results gathered show a high standard deviation in accuracy between participants. In summary, the parameters used with the PSHC processor can mimic the reduced spectral resolution of a CI in a musical context. This could be used to further investigate which parameters could improve the music listening experience for CI users.

7. REFERENCES

[1] B. S. Wilson and M. F. Dorman, "Cochlear implants: A remarkable past and a brilliant future," *Hearing Research*, vol. 242, no. 1, pp. 3–21, 2008, frontiers of auditory prosthesis research: Implications for clinical practice. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378595508001251

[2] T. Goehring, A. Archer-Boyd, J. Arenberg *et al.*, "The effect of increased channel interaction on speech perception with cochlear implants." 2021. [Online]. Available: https://doi.org/10.1038/s41598-021-89932-8

[3] L. Migirov, J. Kronenberg, and Y. Henkin, "Self-reported listening habits and enjoyment of music among adult cochlear implant recipients," *Annals of Otology, Rhinology & Laryngology*, vol. 118, no. 5, pp. 350–355, 2009, pMID: 19548384. [Online]. Available: https://doi.org/10.1177/000348940911800506

[4] J. Marozeau, N. Simon, and H. Innes-Brown, "Cochlear implants can talk but cannot sing in tune," *Acoustics Australia*, vol. 42, no. 2, pp. 131–135, 2014.

[5] S. Paquette, G. Ahmed, M. Goffi-Gomez, A. Hoshino, I. Peretz, and A. Lehmann, "Musical and vocal emotion perception for cochlear implants users," *Hearing Research*, vol. 370, pp. 272–282, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378595518300388

[6] A. Mehta and A. Oxenham, "Vocoder simulations explain complex pitch perception limitations experienced by cochlear implant users," 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5688042/

[7] H. J. McDermott, "Music perception with cochlear implants: A review," *Trends in Amplification*, vol. 8, no. 2, pp. 49–82, 2004, pMID: 15497033. [Online]. Available: https://doi.org/10.1177/108471380400800203

[8] A. Lehmann, C. J. Limb, and J. Marozeau, "Editorial: Music and cochlear implants: Recent developments and continued challenges," *Frontiers in Neuroscience*, vol. 15, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2021.736772

[9] L. K. Holden, J. B. Firszt, R. M. Reeder, R. M. Uchanski, N. Y. Dwyer, and T. A. Holden, "Factors affecting outcomes in cochlear implant recipients implanted with a perimodiolar electrode array located in scala tympani," *Otology and Neurotology*, vol. 37, pp. 1662–1668, 2016. [Online]. Available: https://doi.org/10.1097/MAO.0000000000001241

[10] C. Fuller, D. Başkent, and R. Free, "Early deafened, late implanted cochlear implant users appreciate

music more than and identify music as well as postlingual users," *Frontiers in Neuroscience*, vol. 13, 2019. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2019.01050

[11] C. Karoui, C. James, P. Barone, D. Bakhos, M. Marx, and O. Macherey, "Searching for the sound of a cochlear implant: Evaluation of different vocoder parameters by cochlear implant users with single-sided deafness," *Trends in Hearing*, vol. 23, p. 2331216519866029, 2019, pMID: 31533581. [Online]. Available: https://doi.org/10.1177/2331216519866029

[12] J. J. Galvin-III, Q.-J. Fu, and G. Nogaki, "Melodic contour identification by cochlear implant listeners," 2007. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3627492/

[13] E. D. Casserly, "Effects of real-time cochlear implant simulation on speech production," *The Journal of the Acoustical Society of America*, vol. 137, 2015.

[14] S. A. Ausili, B. Backus, M. J. Agterberg, A. J. van Opstal, and M. M. van Wanrooij, "Sound localization in real-time vocoded cochlear-implant simulations with normal-hearing listeners," *Trends in Hearing*, vol. 23, 2019.

[15] S. F. Poissant, N. A. Whitmal, and R. L. Freyman, "Effects of reverberation and masking on speech intelligibility in cochlear implant simulations," *The Journal of the Acoustical Society of America*, vol. 119, 2006.

[16] N. A. Whitmal, S. F. Poissant, R. L. Freyman, and K. S. Helfer, "Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience," *The Journal of the Acoustical Society of America*, vol. 122, 2007.

[17] S. Jain and P. G. V. Ghosh, "Acoustic simulation of cochlear implant hearing: Effect of manipulating various acoustic parameters on intelligibility of speech," *Cochlear Implants International*, vol. 19, 2018.

[18] J. P. M. Peters, A. W. Wendrich, R. H. M. van Eijl, K. S. Rhebergen, H. Versnel, and W. Grolman, "The sound of a cochlear implant investigated in patients with single-sided deafness and a cochlear implant." *Otology and Neurotology*, vol. 39, pp. 707–714, 2018. [Online]. Available: https://doi.org/10.1097/MAO.0000000000001821

[19] M. F. Dorman, S. C. Natale, A. M. Butts, D. M. Zeitler, and M. L. Carlson, "The sound quality of cochlear implants: Studies with single-sided deaf patients." *Otology and Neurotology*, vol. 38, pp. e268–e273, 2017. [Online]. Available: https://doi.org/10.1097/MAO.0000000000001449

[20] G. Hilkhuysen and O. Macherey, "Optimizing pulse-spreading harmonic complexes to minimize intrinsic modulations after auditory filtering," *The Journal of the Acoustical Society of America*, vol. 136, no. 3, pp. 1281–1294, 2014. [Online]. Available: https://doi.org/10.1121/1.4890642

[21] Q. Mesnildrey, G. Hilkhuysen, and O. Macherey, "Pulse-spreading harmonic complex as an alternative carrier for vocoder simulations of cochlear implants," *The Journal of the Acoustical Society of America*, vol. 139, no. 2, pp. 986–991, 2016. [Online]. Available: https://doi.org/10.1121/1.4941451

[22] Y. Adel, S. Nagel, T. Weissgerber, U. Baumann, and O. Macherey, "Pitch matching in cochlear implant users with single-sided deafness: Effects of electrode position and acoustic stimulus type," *Frontiers in Neuroscience*, vol. 13, 2019. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2019.01119

[23] D. Cavdir, F. Ganis, R. Paisa, P. Williams, and S. Serafin, "Multisensory Integration Design in Music for Cochlear Implant Users," 06 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6798118

[24] R. V. Shannon, F.-G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, "Speech recognition with primarily temporal cues," *Science*, vol. 270, no. 5234, pp. 303–304, 1995. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.270.5234.303

[25] D. D. Greenwood, "A cochlear frequency-position function for several species–29 years later," 2017. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/2373794/

[26] Q. Mesnildrey and O. Macherey, "Simulating the dual-peak excitation pattern produced by bipolar stimulation of a cochlear implant: Effects on speech intelligibility," *Hearing Research*, vol. 319, pp. 32–47, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378595514001841

[27] M. Bernard, "libgammatone: a generic c++ library for gammatone filterbank," 2016. [Online]. Available: https://github.com/mmmaat/libgammatone

[28] J. Galvin, E. Eskridge, S. Oba, and Q. J. Fu, "Melodic contour identification training in cochlear implant users with and without a competing instrument," *Seminars in Hearing*, vol. 33, 2012.

[29] J. J. Galvin, Q. J. Fu, and R. V. Shannon, "Melodic contour identification and music perception by cochlear implant users," vol. 1169, 2009.