SECCIMA: Singing and Ear Training for Children with Cochlear Implants via a Mobile Application

Zhiyan Duan, Chitralekha Gupta, Graham Percival, David Grunberg and Ye Wang
National University of Singapore
{zhiyan,chitralekha}@u.nus.edu, graham@percival-music.ca, {grunberg,wangye}@comp.nus.edu.sg

ABSTRACT

Cochlear implants (CI) can restore part of the hearing of people with severe hearing loss, but these devices are far from perfect. Although this technology allows many users to perceive speech in a quiet room, it is not so successful for music perception. Many public spaces are awash with music, but many CI users do not find music enjoyable or reassuring. This brings multiple challenges to their everyday lives, especially in a world that is saturated with music. Research shows that music training can improve music perception and appreciation for CI users. However, compared to multiple computer-assisted solutions for language training, few such systems for music training are available for children with CI. Existing systems are either targeting a different audience or have complex interfaces that are not suitable for children. In this study, we examined the design limitations of a prior application (MOGAT) and developed a new system with more suitable interfaces. The new system, SECCIMA, was crafted through an iterative design process that involved 16 participants, and the final system was evaluated and compared against MOGAT with another 12 participants. Our results show that SECCIMA is more intuitive and user-friendly than MOGAT.

1 Introduction

Cochlear implants (CI) are electronic devices implanted surgically that directly stimulate the auditory nerve and restore hearing for the deaf. Since their debut in the 1970s, over 120,000 people worldwide have received implants; however, CI are generally optimized for speech rather than music perception [1]. Many CI devices currently in use are limited to 22 electrodes, which each electrode corresponding to one frequency band of vocoder-processed audio. The resulting audio is very dissimilar to how audio is perceived without CIs, and familiar melodies often become completely unrecognizable [1]. Appreciating music can therefore be quite challenging for CI users, who usually have difficulty perceiving both melody and timbre, and do not rank "listening to music" as a pleasurable activity [2, 3]. In a world saturated with music, an inability to appreciate music can dramatically affect one's quality of life [4]. Many adult CI users consider music "noisy" or "disappointing", and listen to music less often than they did prior to hearing loss [4]. However, exposure to music is inevitable in public places. If music is foreign, then one's quality of life in public places will plummet. Greater understanding, even if not complete appreciation of music, will therefore lessen anxiety and lead to a higher quality of life.

Music Training is an effective way to improve music perception and appreciation [4, 5]. However, it is manpower intensive and is conducted periodically, not on a daily basis. Each child can only receive a limited amount of individual attention with current training procedures. It would be beneficial to use mobile devices to help these children engage in learning anytime and anywhere.

In [6], a computer system for adults with CI explored various aspects of music such as timbre, rhythm, mixing and composition. Given its target audience (i.e. adults) and complexity, this system is less than ideal for children. Another work specifically targeted children with CI: MOGAT [7], but while the basic idea behind this work was inspiring, we were concerned that the interface might not be optimal for the target audience. Designing for users — especially children — with disabilities requires extra effort and consideration. Some initial consultations with teachers and therapists gave weight to our concerns, which led to the current project: designing games with user interfaces that are friendly to children with CI.

In this study, we designed a game-based ear training mobile application (SECCIMA) that facilitates the training of musical pitch perception and production skills of children with CI. The game was crafted through an iterative design process that involved 16 children with CI. After the game was finalized, it was compared to MOGAT with an evaluation consisting of 12 new children. We report on our iterative design process, evaluation and comparison results, as well as insights gained while observing the user interaction. Some guidelines on designing user interfaces for children with CI are discussed as well.

The contributions of this work are **i)** designing a game-based music training system that caters to the needs of children with cochlear implants; **ii)** evaluating and comparing the system with a prior application using both objective and subjective measures; and **iii)** summarizing some guidelines for designing user interfaces for children with cochlear implants.

[1] The readers are encouraged to test this themselves with audio such as https://www.youtube.com/watch?v=iwbwhfCWs2Q.

2 Related Work

2.1 Musical Training

Musical training can have a powerful effect on music perception and appreciation. An excellent review of music training for CI users is given in [4]; crucially, it supports the idea that training has the potential to improve music perception and appreciation in this population.

A recent examination of the effect of musical training on CI users found that it improved rhythm, timbre, and melodic contour recognition [8]. These results were echoed in [5], showing that melodic contour identification performance improved sharply after four weeks of music training, and no significant decline was found in the follow-up study eight weeks after the training intervention.

2.2 Computer-aided Training Solutions

There are a number of computer-based **speech** training programs for CI users [9, 10]. However, two recent meta-reviews of computer auditory speech training cautioned that although many studies yielded positive results, the quality of the studies was low to moderate [11, 12]. Other projects explored new possibilities for interaction without performing any usability tests, such as [13], which used an interactive floor to combine movement-based playing and language acquisition , and [14], which allowed children to progress through a story by identifying phonemes.

There are few computer-based **music** training programs. The "Interactive Music Awareness Program" is a computer system for adults which allows users to explore various aspects of music [6], including melodies, rhythms, and texture. A 12-week study with 16 adult CI users showed that IMAP increased their ability to recognize musical instruments. Our main inspiration was "Mobile Game with Auditory Training for Children with Cochlear Implants" (MOGAT) [7]. This project provided three games: "Higher Lower", which prompted children to identify whether a pair of sounds were in ascending or descending pitch; "Vocal Matcher", which asked children to sing notes with specific pitches, and "Ladder Singer", which provided a pitch-graded karaoke interface. However, some flaws in the UI of this system caused problems when being used by children. These are discussed in Section 3.

2.3 Design for Children with Disabilities

In many cases, children with CIs are not as cognitively developed as children with typical hearing, with literacy, language, and behavioral issues due to the cumulative effect of the hearing loss before diagnosis, waiting time for the surgery, and recovery time after the implant [15, 16].

The specific design needs of deaf or hard of hearing children are discussed in two recent papers. [17] noted communication and behavioral difficulties during prototyping sessions, as well as the children being initially hesitant to explore the game being designed. [17] also confirmed that the children were very sensitive to visual changes, quickly noticing (and then fixating on) even minor changes to the background. [18] presented twenty-five guidelines for designing games for deaf children. Many of these guidelines are aimed at mitigating reduced language skills, addressing their sensitivity to visual changes, and mitigating their tendency to be distracted.

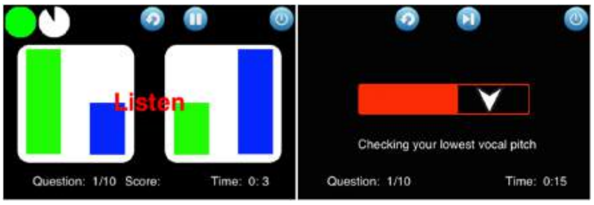


Figure 1: Two interfaces of MOGAT. Left: the "Higher-Lower" game. Right: the "Vocal Matcher" game.

3 Design Considerations

The goal of this study is to design a game-based training application on smartphones that can be used to improve pitch *perception* and *production* skills of children with CIs. Therefore, the application was developed to facilitate the following functions:

1. Play music notes audibly and display them visually;
2. Test players' perception of notes and give feedback;
3. Evaluate players' singing voice, and give feedback on the quality of the pitch produced by the player.

We began by examining MOGAT [7] (Figure 1) to identify design flaws. These are split into the following categories.

3.1 Simplicity

Children with CI are generally less developed in literacy and can be distracted by changing visual elements [17,18]. This means that UI designed for them must be straightforward and simple.

In MOGAT's "Vocal Matcher" game (Figure 1), the child had to first listen to a musical note and try to match that note by singing it. On the screen, there was an empty progress bar that gradually filled up as the child sustained the target pitch. When the pitch produced by the child was different from the target pitch, a small arrow appeared, showing the child how to adjust her pitch to approach the target. However, children sometimes adjusted their pitch in the wrong direction. Later examination revealed that there was an ambiguity in this interface. An arrow pointing down can be interpreted in two ways: either *your pitch is too high; please go lower*, or *compared to the target pitch, yours is too low; please go higher*.

Besides this ambiguity, the arrow design had another potential flaw. Normal progress bars seen in many other interfaces contain no arrows. First, the child was required to notice the arrow, and second, the child must learn the meaning of the arrow. This added to the cognitive load when children were already trying to sing the correct musical note and fill the progress bar.

Finally, screens in MOGAT included a status bar showing the number of completed questions, and a separate timer which counted up. These added little to the actual gameplay, and in the case of the timer could serve to distract the child by inducing them to watch the changing numbers

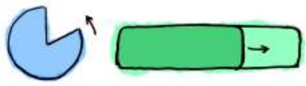


Figure 2: An inconsistency issue in MOGAT.

rather than performing the exercise.

3.2 Consistency/reinforcement

In MOGAT, there were a few inconsistencies between "Higher Lower" and "Vocal Matcher". While the former used a circle that gradually filled itself anti-clockwise during playback, the latter used a horizontal progress bar instead (Figure 2). Similarly, "Vocal Matcher" and "Ladder Singer" both represented musical notes with progress bars, but the latter used the vertical dimension to indicate the player's pitch while the former did not. These small inconsistencies may seem minor, but children with CI were easily confused when presented with such issues. Therefore, a consistent representation of key concepts is necessary.

Although consistency of the meaning of visual elements is important, that meaning does not need to be unique. It would be preferable if we can present information using multiple facets of visual elements without compromising the simplicity of the design.

3.3 Familiarity

Children with CI exhibit a degree of nervousness when facing unfamiliar environments [17]. Therefore, using concrete and familiar objects within the user interface would be a preferable approach when designing for this audience. The visual elements of MOGAT are mostly abstract shapes (e.g. circles, rectangles and bar graphs). While these visual elements look nice with their flat UI style, there were no metaphors between them and real-world objects with which the children are familiar. Therefore, children faced additional challenges in learning the relationship between the abstract graphical representations and the concrete concepts. For example, pitch difference was indicated with bar graphs, which was not taught until Primary 3 according to the local curriculum. Younger children were easily confused because these graphs were alien to them.

4 SECCIMA

In order to introduce familiarity into the new interface, we decided to use two real-world objects as the graphical elements, and attempted to find an analogy for each that was easy for the children with CI to comprehend. One was the *Ball*: when singing, the player's pitch controls the position of a ball. The player needs to drive the ball to a target position by singing the target pitch. The second was a *Balloon*: when singing at a target pitch, the player fills a balloon with air until it bursts; off-pitch singing makes the balloon lose air.

Our approach was to use the concept of reinforcement to exploit different *dimensions* of a single visual element, thus minimizing the cognitive load. Since the core underlying concept of SECCIMA is the pitch of music notes, we need to establish connections between the visual representation of the UI element and the pitch. The visual dimensions we utilized were: *color*, *position* and *movement*.

- *Color*: the red and blue end of the spectrum were used for higher and lower pitches respectively.
- *Position*: higher and lower pitches were positioned near the top and bottom of the screen respectively.
- *Movement*: upward/downward movements were used to represent increasing/decreasing pitch.

Utilizing multiple dimensions of a single object benefits the system in two ways. First, it establishes multiple connections between the target concept and the real-world analogy that children are already familiar with. Multiple connections increase the chances of them understanding the target concept. Moreover, since visual dimensions (color, position and movement) are simply attributes of familiar objects , the connections between these attributes and the object are easy to remember.

The game interfaces in SECCIMA are simple with fewer visual elements compared to MOGAT. We removed multiple indicators like progress bars, arrows and status bars. We used the position of glossy balls to indicate pitch. Children can therefore focus on just moving the ball to the correct position without having to worry about learning multiple indicators. In addition, the appearance and the function of the visual elements used are consistent across all the games. For example, the meaning of color, position, and movement of the glossy balls is the same for all the games.

5 Initial Game Design

There were three games in SECCIMA, namely *Xylophone*, *High/Low* and *Sing'NRoll* (Figure 3). The counterparts of High/Low and Sing'NRoll in MOGAT are Higher Lower and Vocal Matcher. We did not include a karaoke interface in SECCIMA, because the focus of SECCIMA was single pitch perception and production. We built the user interface around the *color*, *position* and *movement* of 8 glossy-looking balls. Each ball is correlated to a musical note with its color and position representing the pitch.

5.1 Xylophone

As the first game in SECCIMA, Xylophone serves to introduce the whole application. In this game, the player simply needs to tap on each glossy ball and listen to the corresponding music note. The purpose of this simple game is to lay the foundation for the rest of the games by helping the players learn the underlying analogies of the UI.

The Xylophone game begins by showing a random subset of the glossy balls. After listening to all notes by tapping all visible balls, the player is taken to the next level. As the levels progress, more balls are shown, ending with all 8 in their respective note positions. This enables to the child to gain familiarity with the interface.

5.2 High/Low

High/Low is the successor of the "Higher Lower" game of MOGAT, and provides training on pitch perception. The interface of High/Low has the same layout as Xylophone, with most of the octave lineup invisible. Two or three

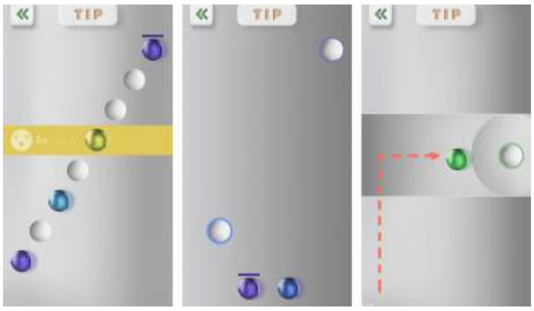


Figure 3: The initial design of the 3 games in SECCIMA: Xylophone, High/Low, and Sing'NRoll *(left to right)*.

glossy balls are placed at the bottom of the screen, while the holes to which those balls belong are shown. The ball's color is the same as in the Xylophone game, while holes are surrounded by a halo of the corresponding color. Upon tapping a ball or a hole, the corresponding musical note is played. To win the game, the player needs to drag the balls to their corresponding holes, ideally by listening to the pitch of each ball and matching it with the correct hole.

As the levels of the game progress, the visual clues gradually disappear so as to challenge the player to rely more on her hearing than on her visual ability. From level three on, the halos around the holes are not colored anymore, and from level six onward, the colors of some balls disappear as well. The final level of this game has all the colors removed, forcing the player to solely rely on her hearing.

5.3 Sing'NRoll

In Sing'NRoll, the player first listens to a musical note (i.e. the target) played by the application, then attempts to reproduce the pitch by singing. The smartphone captures the singing of the player and performs real-time audio analysis to extract the pitch using the YIN algorithm [19]. The user's pitch controls the vertical position of the ball. If the pitch matches the target pitch within a specific tolerance, the ball will be aligned with the hole on screen and start rolling towards the hole, hence the name Sing'NRoll. By sustaining the target pitch for a certain duration, the player can move the ball to the hole and win the game. If the player fails to hold the target pitch, the ball will gradually roll away from the hole, to discourage random guesses.

6 Iterative Design Process

After the prototype was completed, we followed an iterative design process (IDP) to improve the interface. Four rounds of iterative design were conducted, with 3 children in the first, 5 in the second, and 4 in the last two rounds.

6.1 Participants and Procedure

Participants of the user study were recruited from a local primary school dedicated to children with hearing impairment. There were 16 participants in total, with 13 boys and 3 girls. Their age range was 9 to 12 years. All of the participants were pre-lingually deaf, and 14 of them were congenitally deaf. Most of them (14 out of 16) had regular access to smartphones and used smartphones in general to play games. The age range of the participants' cochlear implant was between 10 to 135 months (mean 80.4, std 29.3). All participants had unilateral cochlear implants.

The user studies were conducted in an empty, quiet classroom on the campus of the primary school where the children were recruited. Each session started with a 5-minute briefing, and then the participant played the game for a maximum duration of 15 minutes with no instructions or hints from the test conductors. Finally, a 5-minute feedback conversation was administered and questionnaire filled.

(a) Before (b) After

Figure 4: Sing'NRoll changes after round 1

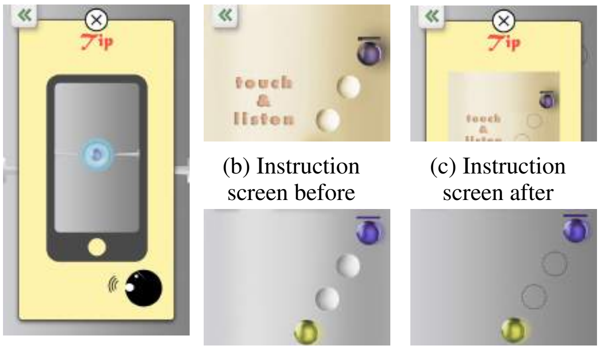
6.2 Round 1

All participants in this round had difficulty figuring out the Sing'NRoll game; they tried dragging the ball rather than singing. This might have been caused by the inconsistency between High/Low and Sing'NRoll, where the action to move a ball in the former was dragging but in the later switched to singing. Another issue was the lack of effective prompt to switch from one action to another. The instruction screen was skipped very quickly by most participants.

Based on these observations, we made the following changes to the design for the next round (Figure 4). For Sing'NRoll, we placed the ball inside a bubble to differentiate it from the balls in the High/Low game. When the player sang, the bubble moved vertically. If the player sustained singing at the target pitch, the bubble gradually grew until reaching some threshold and bursting. By changing the interface dramatically, our aim was to indicate that this game expected different inputs from High/Low. We also removed the hole entirely and replaced it with a target area to avoid any confusion between Sing'NRoll and High/Low. Another consequence of this change is that it avoided a potential confusion regarding movement. The original design used the Y-axis to indicate pitch (which we retain), and the X-axis to indicate completion. By abandoning X-axis movement and using the bubble's size as a completion metric, we clarify the meaning of movement in the game. Finally, a prompt message was added to remind the player to sing if she remained silent for more than 6 seconds.

6.3 Round 2

Neither the bubble nor the prompt message seemed to have any effect. Only two out of the five participants managed to figure out how to play the game. Although the prompt said explicitly "*Sing to move bubble. Break the bubble here to get the Gem.*", children did not seem to follow it. Young children with CI have difficulty understanding long



(a) Animated tip of Sing'NRoll (b) Instruction screen before (c) Instruction screen after (d) Xylophone before (e) Xylophone after

Figure 5: UI changes for round 2

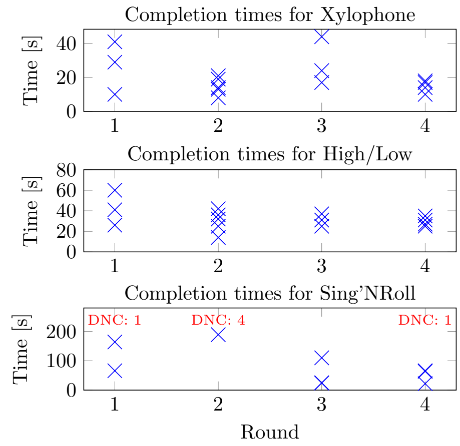


Figure 6: Completion times for games. "DNC" indicates that the child "did not complete" the exercise.

sentences [17, 18], and we failed to adequately heed those warnings. The arrow pointing at the target area encouraged participants to tap on that area, which caused more confusion. A few participants also tried to play the game directly on the instruction screen, because it had a very similar look to the actual game and caused the confusion.

Another subtle consistency issue was observed. Since holes in Xylophone were not tappable, participants considered holes in High/Low to behave the same way. However, the later were designed to play the note upon tapping.

We made three changes after this round: 1) instruction screen of Sing'NRoll was changed from a static image to an animated demo (Figure 5a), 2) all instruction screens received a style update different from the actual games (Figure 5b, 5c), 3) holes in Xylophone and High/Low were given different appearances (Figure 5d, 5e).

6.4 Round 3 & Round 4

The animated cartoon instruction worked well. All four participants learned to play the Sing'NRoll game very quickly (three of them managed to finish the first level in under 30 seconds). The other two problems observed in round 2 were also resolved by the corresponding adjustments.

With no obvious flaws found in round 3, we added a scoreboard to Sing'NRoll as an extra motivational component to see if it can boost excitement.

No new problems were discovered in round 4, but there was also no effect of the scoreboard — all of the children quickly skipped over it. The scoreboard might contain more information than they would like to consume.

6.5 Objective Evaluation

During the iterative design process, participants' interactions with SECCIMA were filmed, from which we calculated the time taken for a participant to complete each game and used this time as an objective measure to quantify the intuitiveness and ease of use of the UI.

For Xylophone and Sing'NRoll, the completion time was defined as the time from a player entering the game to winning the first level of the game. The completion time of the High/Low game was defined differently because the first level of High/Low has color strokes around the targets and the participant can win the game by simple color matching. Therefore we defined the completion time of the High/Low game as the time lapsed from the player entering the game to the player winning the third level, where there are no color indications around the targets.

No.	Statement
Q1	I understood the instructions
Q2	The game was fun
Q3	The game was easy to play
Q4	I want to play this game again
Q5	I want to play this game at home
G1	Rate the Xylophone game
G2	Rate the High/Low game
G3	Rate the Sing'NRoll game

Table 1: Post-session questionnaire for participants.

Figure 6 shows the completion time of all participants organized by round number and game. One participant in round 3 was excluded because the corresponding video recording was corrupted. Both the Xylophone game and the High/Low game were completed within a minute (many within 30 seconds). The completion time for the Sing'NRoll game was more interesting as this was the challenging part for the participants. In the first two rounds when there were no animated instructions, 5 out of 8 participants failed to complete the game. Those who managed to complete the game had struggled for quite some time (89, 164, 189 seconds) to succeed. After the introduction of the animated instruction in round three, the majority of the participants (7 out of 8) completed the game without much struggle. The completion time was reduced significantly as well, from more than two minutes to well under one minute.

6.6 Subjective Evaluation

Subjective measurements were collected via the post-session questionnaire (Table 1, Figure 7). The participants were asked to choose from a Likert scale of 5, with larger numbers corresponding to more positive answers. The responses to Q1 (*I understood the instructions*) and Q3 (*The game was easy to play*) exhibit a high correlation to the introduction of animated instruction, as they increase noticeably after round 2. This confirms our observation from the objective evaluation that the animated instruction did help the participants learn the game faster and make the

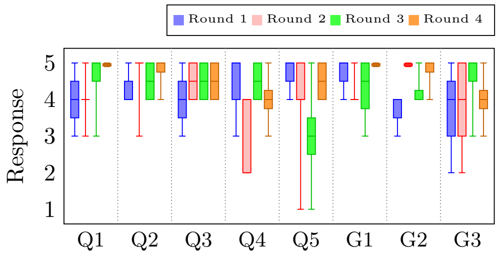


Figure 7: Questionnaire results for each round (box plots).

game easier. The scores of Q2 (*The game was fun*) increased marginally over time; however, the large variations in scores of Q4 and Q5 suggest that these improvements in enjoyment were not significant.

The participants were also asked to rate the three games on a Likert scale of 1 to 5. It can be seen that the scores of Xylophone (G1) and High/Low (G2) are noticeably higher than that of Sing'NRoll (G3). One possible reason is that in the Sing'NRoll game participants were facing more challenges, which may have attenuated their enjoyment. But even though Sing'NRoll is less preferred, its scores are on a upward trend, which points to the effectiveness of the adjustments to the system.

7 Comparison with MOGAT

SECCIMA was compared with MOGAT to assess its intuitiveness. In order to distinguish the two, we will prepend the game names with S- or M-, respectively. Both applications have two main components: a *listening game* (LG: M-Higher Lower vs S-High/Low) and a *singing game* (SG: M-Vocal Matcher vs S-Sing'NRoll in SECCIMA).

7.1 Participants, Setup and Procedure

A total of 12 participants were involved, out of which 7 were congenitally deaf, 4 pre-lingually deaf and 1 post-lingually deaf. The age range was 7 to 10 years. All participants were regular smartphone users and had used them for playing games. Time since cochlear implantation ranged from 20 to 84 months (mean 45.3, std 22.9).

For SECCIMA, the same equipment as in the iterative design process was used. For MOGAT, we used an iPod Touch (4^{th} gen.). The same audio cable was used to connect the device with participants' implants.

All participants had **no prior exposure** to either interface. They were randomly divided into two groups, each using one of the apps. The same procedure as in the iterative design process was used, i.e. a pre-session briefing, 15 minutes of game play (video recorded), and a short discussion as well as a post-session questionnaire.

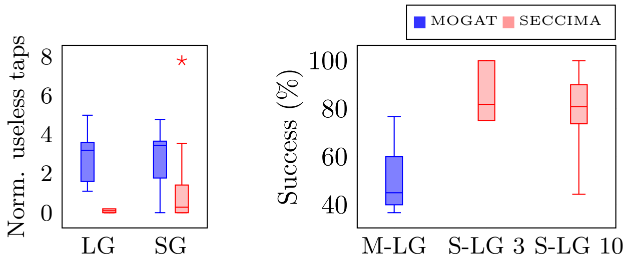
7.2 Objective Evaluation

Objective measures were derived from the video recordings. Since one participant opted out of video recording, we were left with 6 subjects in the SECCIMA group and 5 in the MOGAT group.

We counted the *unnecessary taps* (i.e. tapping on UI elements that are not tappable) by the participants in both groups (shown in Table 2), which was then divided by the total number of UI elements in respective games to normalize the difference in relative densities of screen layouts. We used this normalized number of unnecessary taps as an objective measure of the intuitiveness of the UI, with more unnecessary taps indicating a poor interface. As shown in Figure 8a, the normalized number of unnecessary taps in SECCIMA were significantly fewer than MOGAT in the listening game. Results for the singing game also favored SECCIMA, but the difference was not significant.

App	No. of UI elements	No. of useless taps
M-LG	10	11, 16, 32, 36, 50
S-LG	6	0, 0, 0, 0, 0, 0
M-SG	9	0, 16, 31, 33, 43
S-SG	7	0, 0, 0, 4, 10, 55

Table 2: Number of unnecessary taps from users



(a) Normalized unnecessary taps from users (*lower is better*). $p = \mathbf{0.001}, 0.51$.

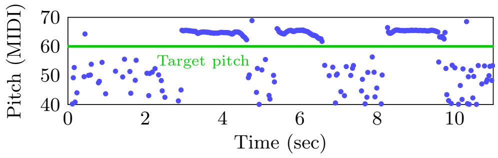
(b) Success rate in listening games (*higher is better*), S-H/L x: first x levels of S-High/Low. $p = \mathbf{0.006}, 0.06$.

Figure 8: Objective comparison (box plots).

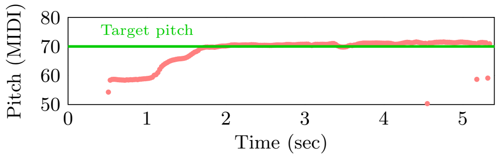
Another objective measure we used was the percentage of successful attempts in the listening game. In both interfaces, the child was supposed to distinguish musical notes and to indicate which one is higher or lower. Choosing randomly would give a 50% successful rate. Figure 8b shows the successful rates for M-Higher Lower and S-High/Low. With a mean success rate of 50.67% (vs. 81.8 % and 80.8 % in SECCIMA), children playing M-Higher Lower were likely to be making selections randomly, indicating the user interface was not intuitive enough for the children to understand. Later levels of S-High/Low were more challenging because they involved more than two musical notes (only the first three levels involved two notes), but as student performance remained high, the interface of SECCIMA was demonstrated to be more effective.

In the singing game, **none** of the 6 children using MOGAT could figure out how to play the game on their own until the teacher intervened. On the contrary, 3 out of 6 children in the SECCIMA group managed to figure it out on their own, and another one kept blowing air instead of singing. Only two of them failed to understand that they were supposed to make sound. Although not as good as in the iterative design study (in which all but one child in rounds 3 and 4 figured it out), this result was not surprising considering the younger age of this group (7–10 vs. 9–12).

In terms of feedback, the children's pitches in the MOGAT group were mostly monotonic, suggesting that they did not fully understand the feedback. In the SECCIMA group, more children were varying their pitches in response



(a) One example of a pitch contour from MOGAT



(b) One example of a pitch contour from SECCIMA

Figure 9: Comparison of two example pitch contours.

to the feedback. As an example, Figure 9a and 9b show two typical pitch contours from users of MOGAT and SECCIMA respectively. The pitch contour from MOGAT was mostly monotonic and did not move towards the target pitch over time. On the other hand, the one from SECCIMA slid nicely towards the target pitch and stayed there.

For the comparison between the two *singing* games, "completion time" was not used as a comparing metric because the game designs were too different to compare fairly. For example,M-Vocal Matcher did not penalize the user for incorrect pitch, which made it much easier than S-Sing'NRoll, because the latter imposed a penalty for incorrect pitch and therefore was harder to win. For the same reason, success rate was not used either.

7.3 Subjective Evaluation

Subjective measurements were collected via the post-session questionnaire (questions in Table 1, results in Figure 10).

The children indicated that SECCIMA was easier to understand and play than MOGAT (Q1 and Q3). It is worth noting that some of the children playing MOGAT indicated they understood the game fully (by rating Q1 with a score of 5), although they did **not** play the game as they were supposed to. In fact, none of the children in the MOGAT group was able to figure out how to play the singing game (Vocal Matcher) after a long struggle, and the teacher had to intervene. Discussions with their teacher confirmed that they tend to say *Yes* to those questions they do not understand. This could potentially bring inaccuracies to the results. Researchers working on comparisons between people with different ages or mental capabilities may need to take this potential bias into consideration.

The children's teacher who helped us conduct the study provided some feedback on both interfaces. She mentioned that MOGAT was less intuitive. For example, the word *Select* used in MOGAT was not in the children's vocabulary, and the children had no exposure to bar graphs until their later stage in primary school. In contrast, SECCIMA's colored balls were simpler and easier to understand. She specifically pointed out that the arrows (see Figure 1) used by Vocal Matcher in MOGAT were confusing, even to her as an adult. Children were tapping on the flashing arrow even though it was not tappable. On the other hand, the animation used in SECCIMA was more intuitive.

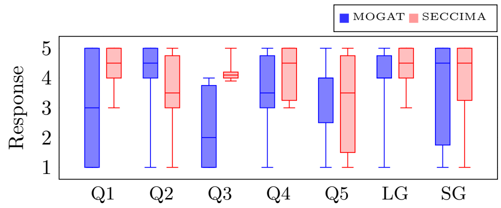


Figure 10: Questionnaire results (*higher is better*), see text for details about possible inaccuracies (box plots). $p = 0.21, 0.58, \mathbf{0.03}, 0.39, 0.74, 0.49, 0.76$.

8 Discussion

The interfaces for SECCIMA games were simpler than MOGAT, and the visual elements were consistent throughout all the games. The lower number of unnecessary taps and the higher success rate in SECCIMA confirms our hypothesis that increasing the simplicity and the consistency of the games made them easy to follow. Additionally, the reinforcement considerations resulted in a step-by-step learning process in SECCIMA. The lesson about color and position of the balls in the first game helped the students understand how to drag the balls to correct positions in the second game. Similarly, lessons from each level were being applied to play the next levels. The reduction in unnecessary taps and more-than-chance success rate of SECCIMA shows that the children were learning the games, and not randomly playing them. Finally, in SingN'Roll, familiarity with floating bubbles may have helped the children to understand the feedback on how to change their pitch. Thus, it was a compound effect of all design considerations that made the interface more intuitive and less confusing.

We offer a few recommendations for researchers working with children with CI. **(1) Test early, test often**: No matter how many hours are spent discussing designs in the lab, something unexpected will always happen when testing with children. That said, it is important to **(2) carefully balance the number of children in the design vs. evaluation phases**. There are usually limited number of CI users in each city, and parents and caregivers are appropriately cautious about children's participation in research studies.

It might be helpful to **(3) involve participants outside of the target population**. While the real evaluation should still be done with the target group, some of the flaws could have been detected by children with normal hearing.

It is important to **(4) reduce the amount of text and use more visual displays**, provided those visuals displays do not cause distractions. **(5) When text is necessary, words and phrases are more preferable then sentences**, because these children have difficulties comprehending longer text. It is worth noting that even the use of imagery and animation should be controlled to minimize cognitive load. These findings are consistent with those of [17, 18].

We observed that many young children with CI are not able to express themselves effectively (also noted in [16, 17]); forming sentences can be difficult for them. Therefore, **(6) obtaining verbal feedback from them takes time and patience** (as with adults with cognitive disabilities [20]). As noted earlier in the comparison study, some children indicated that they completely understood the instructions while they actually did not. When collecting subjective feedback from these children, this "*nodding phenomenon*" should be taken into consideration to avoid potential bias. We therefore suggest to **(7) use objective measurements in addition to subjective feedback**, such as success rate, number of unnecessary taps, and completion times.

9 Conclusion

We designed SECCIMA to facilitate music training of children with cochlear implants using an iterative design process. The observations and design decisions in each iteration were documented. Both objective and subjective evaluations were performed to assess the effectiveness of the design decisions and the ease of use of the interface. The evaluation result shows that the adjustments between iterations improved the usability and made the interface easier to use. A comparison with an existing work was performed and showed that our system is more intuitive and easier for the children to use. The insights and guidelines summarized from this study may be helpful to those who are interested in designing applications for children with CI.

Acknowledgments

The authors would like to thank Ms. Joan Tan for facilitating the user study as well as many excellent suggestions to our project. We are grateful to Ms. Terry Theseira and her colleagues and students from Canossian School for assistance and participation in the user studies. This project was partially funded by a research grant R-252-000-597-281 from National Research Foundation in Singapore.

10 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.

[2] H. J. McDermott, "Music perception with cochlear implants: a review," *Trends in amplification*, vol. 8, no. 2, pp. 49–82, 2004.

[3] C. J. Limb and A. T. Roy, "Technological, biological, and acoustical constraints to music perception in cochlear implant users," *Hearing Research*, vol. 308, pp. 13 – 26, 2014.

[4] V. Looi, K. Gfeller, and V. Driscoll, "Music appreciation and training for cochlear implant recipients: a review," in *Seminars in hearing*, vol. 33, no. 4, 2012, p. 307.

[5] Q.-J. Fu, J. J. Galvin, X. Wang, and J.-L. Wu, "Benefits of music training in mandarin-speaking pediatric cochlear implant users," *Journal of Speech, Language, and Hearing Research*, pp. 1–7, 2014.

[6] R. M. van Besouw, B. R. Oliver, M. L. Grasmeder, S. M. Hodkinson, and H. Solheim, "Evaluation of an interactive music awareness program for cochlear implant recipients," *Music Perception: An Interdisciplinary Journal*, vol. 33, no. 4, pp. 493–508, 2016.

[7] Y. Zhou, K. C. Sim, P. Tan, and Y. Wang, "Mogat: mobile games with auditory training for children with cochlear implants," in *ACM Multimedia*. ACM, 2012, pp. 429–438.

[8] B. Petersen, M. V. Mortensen, M. Hansen, and P. Vuust, "Singing in the key of life: A study on effects of musical ear training after cochlear implantation." *Psychomusicology: Music, Mind, and Brain*, vol. 22, no. 2, p. 134, 2012.

[9] Q.-J. Fu and J. J. Galvin III, "Maximizing cochlear implant patients' performance with advanced speech training procedures," *Hearing research*, vol. 242, no. 1, pp. 198–208, 2008.

[10] ——, "Computer-assisted speech training for cochlear implant patients: Feasibility, outcomes, and future directions," in *Seminars in hearing*, vol. 28, no. 2. NIH Public Access, 2007.

[11] H. Henshaw and M. A. Ferguson, "Efficacy of individual computer-based auditory training for people with hearing loss: A systematic review of the evidence," *PLoS ONE*, vol. 8, no. 5, p. e62836, 05 2013.

[12] R. Pizarek, V. Shafiro, and P. McCarthy, "Effect of computerized auditory training on speech perception of adults with hearing impairment," *SIG 7 Perspectives on Aural Rehabilitation and Its Instrumentation*, vol. 20, no. 3, pp. 91–106, 2013.

[13] O. S. Iversen, K. J. Kortbek, K. R. Nielsen, and L. Aagaard, "Stepstone: An interactive floor application for hearing impaired children with a cochlear implant," in *Interaction Design and Children*, ser. IDC '07. New York, NY, USA: ACM, 2007, pp. 117–124. [Online]. Available: http://doi.acm.org/10.1145/1297277.1297301

[14] S. Cano, V. Peñeñory, C. A. Collazos, H. M. Fardoun, and D. M. Alghazzawi, "Training with Phonak: Serious Game As Support in Auditory – Verbal Therapy for Children with Cochlear Implants," in *ICTs for Improving Patients Rehabilitation Research Techniques*, ser. REHAB '15. New York, NY, USA: ACM, 2015, pp. 22–25. [Online]. Available: http://doi.acm.org/10.1145/2838944.2838950

[15] R. Calderon and M. Greenberg, "Social and emotional development of deaf children: Family, school, and program effects," *The Oxford handbook of deaf studies, language, and education*, vol. 1, pp. 188–199, 2011.

[16] S. C. Theunissen, C. Rieffe, M. Kouwenberg, L. J. De Raeve, W. Soede, J. J. Briaire, and J. H. Frijns, "Behavioral problems in school-aged hearing-impaired children: the influence of sociodemographic, linguistic, and medical factors," *European child & adolescent psychiatry*, vol. 23, no. 4, pp. 187–196, 2014.

[17] L. E. Potter, J. Korte, and S. Nielsen, "Design with the deaf: Do deaf children need their own approach when designing technology?" in *Interaction Design and Children*, ser. IDC '14. New York, NY, USA: ACM, 2014, pp. 249–252.

[18] A. Melonio and R. Gennari, "How to design games for deaf children: Evidence-based guidelines," in *2nd International Workshop on Evidence-based Technology Enhanced Learning*, ser. Advances in Intelligent Systems and Computing. Springer International Publishing, 2013, vol. 218, pp. 83–92.

[19] A. De Cheveigné and H. Kawahara, "Yin, a fundamental frequency estimator for speech and music," *J. of the Acoust. Soc. of America*, vol. 111, no. 4, pp. 1917–1930, 2002.

[20] S. Johansson, J. Gulliksen, and A. Lantz, "User participation when users have mental and cognitive disabilities," in *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*, ser. ASSETS '15. New York, NY, USA: ACM, 2015, pp. 69–76. [Online]. Available: http://doi.acm.org/10.1145/2700648.2809849