Motivating Individuals with Spastic Cerebral Palsy to Speak Using Mobile Speech Recognition

Zak Rubin, Sri Kurniawan, Taylor Gotfrid, Annie Pugliese
Dept. of Computer Engineering, Jack Baskin School of Engineering
University of California Santa Cruz, 1156 High St, Santa Cruz, CA 95060, USA
Ph: +1-831-459-1037
Email: {zarubin, skurnia, tgotfrid, abpugliese}@ucsc.edu

ABSTRACT
Individuals with cerebral palsy (CP) struggle with conditions such as dysarthria, dysphagia, and dyspraxia as they speak. While speech therapy is successful in practice, outside practice requires increased commitment and effort from caregivers. Researchers developed a speech recognition game designed to encourage out-of-office exercises and motivate users to practice. Next they recruited a participant with cerebral palsy to investigate the performance of the system in a live environment. The participant joined the game after demonstration from the caregiver and temporarily increased speech loudness and clarity during play. The participant found sound effects more rewarding than animations. The total number of sentences spoken during the session was found to be less than half that of a speaker without any impairment. Researchers also observed two instances of cheating. This work provides insight into the automated motivation of motivating speech production with individuals with cerebral palsy.

CCS Concepts
• Human-centered computing~Usability testing abase management system engines

Keywords
Cerebral Palsy; Speech Recognition Interface; Speech Therapy; User Study

1. INTRODUCTION
More than half of all individuals with cerebral palsy deal with some form of speech disorder [1]. These speech problems result in reduced ability to function in social situations leading to social isolation and depression. Cerebral palsy speech is considered dysarthric and is characterized by a variety of speech issues including slurring, chopping, speed issues, mouth movement issues, and voice quality issues [2].

Treatment of dysarthria presently involves direct intervention by a speech-language pathologist. Two-thirds of those who attempt speech therapy improve their intelligibility and communication functioning. They also became intelligible to all listeners following intervention. The speech language pathologist determines if assistive technology would benefit the individual [3].

Despite assistive technology's potential benefit to this realm research conducted in 2013 found that speech language pathologists were unaware of any available assistive technology for those with speech impairments. An investigation of existing technologies found no technology developed to interact with speech impairments. Interviews with the primary stakeholders in speech therapy determined that a tool that motivated individuals in therapy and provided speech therapists with feedback about their practice and performance outside of the office would provide the most impact in the lives of those with speech impairments [4].

ASSETS '16, October 23-26, 2016, Reno, NV, USA
ACM 978-1-4503-4124-0/16/10.
http://dx.doi.org/10.1145/2982142.2982203

2. RELATED WORK
Researchers at the University of California, Berkeley investigated computerized speech recognition of individuals with dysarthria caused by cerebral palsy. Nine athetoid cerebral palsy speakers and one spastic speaker took part in a test investigating the performance of the Shadow/VET voice entry system. Researchers found that dysarthric speakers had a similar pattern of correct recognition to non-disabled speakers but the correct recognitions of dysarthric speakers was less than half that of non-disabled speakers. The researchers noted that the low recognition accuracy remains a serious problem [5].

Speech Adventure is a speech therapy tool suite intended for use by children enrolled in speech therapy. Speech Adventure utilizes a speech recognition engine capable of detecting coarse mispronunciations characteristic of speech impairments related to cleft lip and palate [6]. Researchers adapted the engine used in Speech Adventure to ignore mispronunciations and developed games focused on sound effects and animations.

3. SPEECH GAME
A graduate and undergraduate student developed a total of two levels. Five levels intended for plosive speech practice had their dictionaries replaced with command and control dictionaries containing fewer than ten words. Words were translated into their Arpabet equivalent via the lmtool provided by Carnegie-Mellon University [7].

These dictionaries were put into a rotating system of a total of seven mini-games. The game engine cycles through the games randomly, making sure each game is played at least once before repeating. Each level's time was set to 20 seconds, if the user does not finish the level in the allotted time the game moves to the next level. The system provides a timer for each level as well as on-screen prompts to guide the user what to say. As the user speaks the on-screen phrase highlights as the user speaks. If the user

speaks the word or phrase correctly, an action occurs such as a sound effect or an animation.

Table 1. Sample In-Game Dictionary

Word	Arpabet Translation
BLAST	B L AE S T
BOOM	B UW M
BURST	B ER S T

Researchers evaluated the system on four native English speakers with no speech disorders. Three females and one male aged 21-27 took part in a playtesting session lasting a half hour. The mean speech rate in game was 17.2 words (SD = 1.92) per minute of playtime.

4. USER STUDY

Undergraduate students from the University of California, Santa Cruz recruited a 34-year-old female participant with spastic cerebral palsy as part of a course introducing beneficial technology to disabled groups. The participant had previous experience playing a variety of games on her own tablet. Testing took place in a quiet indoor location on campus. The participant's home support was present in place of the primary caregiver.

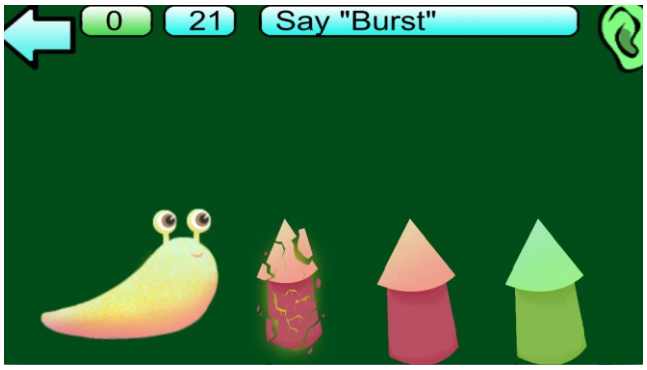

Figure 1 Fireworks Level. Players touch a rocket and then say the word in the text box to launch the rocket.

Initially the participant was reluctant to participate. The home support motivated the participant by playing tutorial levels. On the second in-game sentence the participant attempted to play, but let the home support finish the sentence. The participant began playing the game on her own on the second level of the tutorial and played along with the home support for the remainder of the levels as they played the main game together. The participant reacted positively to all in-game sound effects. During play of the main game (levels 3-10), the home support and observers noted a significant increase in the participant's pronunciation and volume.

Researchers witnessed the participant cheating in the game on two separate occasions. Cheating in this context was defined as any attempt to gain points or progress in the game without saying the on-screen sentence. In the first instance the participant exploited the caregiver's verbal motivation for all of game level seven. The participant would attempt to have the caregiver say the word loud enough so that the recognition engine would respond to the caregiver. In the second instance the participant could sometimes trigger events in the game through any vocalization. This began on level eleven and continued until the game concluded.

The game reported that the participant played for 4 minutes. During that time they played 12 levels and spoke 34 valid sentences, for a speech rate of 8.5 sentences per minute. Of those 8 were utterances from the participant; the remainder came from the home support. Nine of the twelve levels played were unique and three were repeats. The participant expressed disinterest in the game at level ten, and was unable to complete any level before time ran out. The home support reported that the participant enjoyed the game overall, but grew bored primarily due to the repetition of the games.

5. DISCUSSION

We have described a speech recognition tool suite designed to motivate individuals with speech disorders to practice their voice in an engaging virtual environment. Testing on an adult participant with spastic cerebral palsy resulted in a temporary increase in both speech loudness and pronunciation. The system's initial results are promising in improving speech characteristics of those with speech impairments as well as the development of speech recognition systems designed for disabled users.

The adult participant required significant motivation from home support. The system also did not maintain the individual's interest beyond the first cycle of mini-games. Finally, the participant resumed normal speech patterns shortly after use so more application is necessary to realize the benefits. It is difficult to determine the degree of reluctance of the participant to join initially as the study was conducted in an unfamiliar environment. This work would most benefit from an expansion into home environments for long-term investigations to determine if the improvements transfer over to real-life conversational scenarios.

6. REFERENCES

[1] Nordberg, A., Miniscalco, C., Lohmander, A. and Himmelmann, K., 2013. *Speech problems affect more than one in two children with cerebral palsy: Swedish population-based study.* Acta paediatrica, *102*(2), pp.161-166.

[2] *Dysarthria*. ASHA http://www.asha.org/public/speech/disorders/dysarthria/ Retrieved 6/19/2016.

[3] Yorkston, K. *Treatment Efficacy Summary: Dysarthria (Neurlogical Motor Speech Impairment)*. ASHA, 2003.

[4] Rubin, Z., Kurniawan S. *Speech Adventure: Speech Recognition for Cleft Speech Therapy.* Proceedings of 6th International Conference on Pervasive Technologies Related to Assistive Environments. PETRA 2013.

[5] Coleman, C., Meyers, L. *Computer Recognition of the Speech of Adults with Cerebral Palsy and Dysarthria*. Augmentative and Alternative Communication, 7:1, 34-42. 1991. DOI: 10.1080/07434619112331275663

[6] Rubin, Z., Tollefson, T., Kurniawan S. 2014. *Results from Using Automatic Speech Recognition in Cleft Speech Therapy with Children.* Computers Helping People with Special Needs, 283-286.

[7] Lenzo, K. *The CMU Pronunciation Dictionary*. http://www.speech.cs.cmu.edu/cgi-bin/cmudict. Retrieved 6/19/2016.