Early Language Assistance Therapy (ELAT) Robot for Children with Autism Spectrum Disorder (ASD)
Maurie Zhang, Myles Lewis, Eugene Hsiao, Andrew Yoon, Miguel Esteban, Jonathan Ko, Remi Shittu

Summary and Specific Aims

The goal of this project is to help ASD affected children improve their language skills and build their language intelligibility to a level to where they can practice and interact more with other children and people. The choice of the dog CHiP WowWee robot helps the child feel at ease and makes learning a fun experience. The project will integrate the team-created AI with the WowWee SDK to create a language therapy curriculum that will supplement speech therapy sessions, and allow for practice from anywhere[1].

Specific Aim 1: Develop a software to detect vocalizations, calculate a percent similarity value, and adapt.

The purpose of the software is to allow for sound funneling, to help the child refine, incoherent vocalizations, into more intelligible sounds, understandable to all English speakers. The child's iteration of the model word will be compared to accepted pronunciations of the model world; the software will return a similarity index as a percentage of closeness to the accurate pronunciation. The software will be adaptable, i.e. if the concepts are too easy the program will move on to harder concepts, and vice versa.

Specific Aim 2: Integrate the developed software into an affordable robot.

In order to make language therapy more accessible it must be affordable, so the team will integrate the developed AI Software with existing WowWee robot CHiP, which retails for $100[2].

Specific Aim 3: Provide a language therapy curriculum

The programmed robot package will provide an expressive language curriculum and allow for dynamic interaction, allowing meaningful, interactive and self-paced practice outside of traditional clinical therapy sessions. With 200 words and phrases in the database, the robot will start on a low level and help the child practice with simple words. Eg. Mom, dad, dog, cat, etc. The curriculum will be cumulative and so the child will be quizzed with concepts from all past levels to cement understanding.

The project aims to help ASD impacted children improve their language and social skills, make them feel more confident with their capabilities, so they can reach their full potential.

Research/Design Strategy

A. Significance

Autism Spectrum Disorder (ASD) impacts the lives of children (and adults) of all racial, ethnic, and socioeconomic groups, the disorder is described as a spectrum, in that it does not have the same effects on all people, however, in general children with ASD are language impaired in comparison to their non-ASD-impacted peers, and it is this initial language setback which leads to developmental delays and significant impairments to cognitive and social skills[3]. Research from the Center of Disease Control (CDC) shows that ASD affected children have great untapped potential, "… almost half (44%) of children identified with ASD has average to above average intellectual ability"[4]. A language deficiency also makes it difficult for an individual to understand him/herself well, because the person has difficulty understanding relationships. Stronger language skills allow for children " [to] have better self-control" and self-confidence[5]. The project proposal aims to create a more accessible therapy solution which will help the ASD affected children feel encouraged and confident in practicing vocalizations, build adaptive skills, so they can reach their full potential.

Current therapy options rely heavily on specialized health practitioners, mainly Speech Language Pathologists (SLP), but also ASD-trained professionals, behavioral and occupational therapists, all of whom are highly concentrated in urban areas, and in general, not accessible to the majority of the children with ASD, for logistic (scheduling, location) and financial reasons[6]. Insurance companies and other third-party payers are interested in investing in "effective and necessary treatments" and there are some insurance plans with no coverage for Applied Behavior Analysis (ABA) Therapy, speech therapy is an ABA therapy[78].

Implementation of this project, will allow for base-line diagnostics and language therapy practice to be performed from anywhere, allowing for SLPs and other clinicians to develop more personalized and effective therapy for each child. Implementation of a functional technology described earlier in the Specific Aims section, if integrated with traditional clinic practices, could lead to more insurance covered treatments related to speech therapy.

B. Innovation

This project proposal in ways supports the roles of traditional speech and therapy clinicians, and yet also challenges their current roles, in that the technology implemented would be able to produce base-line diagnostics and practice with the child, which would require therapy to be more engaging and provide more rigorous practice.

All the components we intend to use already exist (blue-tooth, WowWee Robot, JAVA, Tensor Flow), so the challenge is not designing new components, but rather in reconfiguring the network and sub-systems to optimize the potential of each component. The noveler part of the project would be in developing an effective child vocalization classification/grading system. Adult vocalization libraries exist, however they work effectively only for adult vocalizations. This is because child vocal cords are not fully developed so even when examining the vocalizations of an individual child for a single word or vowel there is a greater range of fluctuation in the pronunciation in comparison to the range of pronunciations an adult would produce, i.e. the child categorization system needs to consider more variables. To address this difference, the group will be creating a new classification system based on a library of only child vocalizations. The AI subgroup will look into applying traditional linguistics (phoneme analysis) and digital signal processing(DSP) techniques as well as developing an algorithm with Tensor Flow to develop a more efficient search and match process. Tensor flow is, "an open-source software library for dataflow programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks[9]." Researchers and industry related Tensor

Flow projects (diagnostics, image processing, music production) have proven that the Tensor Flow has great potential to sort through complex multidimensional data sets to give clearer results and images[10]. The approach will be to test both the traditional and new (Tensor Flow, Formant Analysis[11]) in parallel with traditional (DSP, Phoneme) analysis methods, so when the robot-software team has finished their work, the team can implement the best sorting algorithm available.

C. Approach

Aim 1: Develop a software to detect vocalizations, calculate a percent similarity value, and adapt.
The purpose of the software is to combine the sound categorization and grading into an algorithm (percent similarity index) with an encouraging nature to make the child feel comfortable practicing skills.

Aim *1.1. Design a program that encourages the child to produce sounds that are of 50% or greater intelligibility.*
There will be 5 levels to the program. Each level except for the first level will grade the vocalization and only lead the child onto a higher level if the vocalization attempts are greater or just as accurate as the baseline percent similarity index. Level 2 encourages the child to vocalize the word at 10% or greater accuracy. Each higher level increments 10% more in base-line percent similarity, and ultimately the goal of this program is to get the child to produce words/phrases that are 50% (or greater) intelligibility, 50% of the words or phrases are understood by other English speakers.

Each level gives the child 5 attempts to vocalize the word. The robot will give the accurate pronunciation and encourage the child to repeat each time. The robot will also repeat the word/phrase at the end of each level to reinforce the concept.

Level 1: The robot will start the interaction by saying the model word/phrase and then give the child a preview of the reward (a dance or a song). The robot will then prompt the child to repeat the word/phrase and will dance for any vocalization (accurate or not) and move onto the next level after 5 vocalizations. The goal is to encourage the child to talk and feel comfortable interacting with the robot CHiP. *Level 2:* At the second level, this is when the program will start analyzing the vocalizations for similarity percentages. If the sound is greater than 10% similarity the robot will reward the child and the child will move onto Level 3. If the child is unsuccessful and produces “bbbbbbrr” instead of “aaaaafff” for a word that has a beginning “aff” sound, then the program will grade that as incorrect and drop down to Level 1 and then move onto Level 2 again to try again with the same word/phrase. Level 3 – Level 5. Are essentially repeats of Level 2, but with a higher base-line percentage accepted.

Challenge: Program needs to be adaptable and provide a way for the child to improve.
The goal of the program is to provide a way for the child to practice meaningfully. There is a possibility that the child will have trouble with certain concepts and be stuck on a single concept or a string of concepts and get stuck there for a long time. The laid out design considers this by moving the child onto a lower level for a brief period and then moving up back to the higher level again to try a similar concept, that perhaps the child is better with. Working on a similar sound set will allow the child to practice and then when they return to the first missed concept they will be able to master that as well.

Aim 2: Integrate the developed AI software into an affordable robot.
The team will integrate the developed AI Software with existing WowWee robot, the CHiP model[12], which retails for $100 and build on previous groups’ work by using Blue-Tooth technology to create a secure and stable connection between the robot and the android device.

Aim 2.1: With a library of only child vocalizations create a categorization system that can process child sounds accurately.
There exist standard speech processing libraries for adult sounds, however, these can not be used for analyzing child vocalizations, because child vocalizations are not as precise and have a greater range of fluxuations, adult sound processing libraries would not be able to process child vocalizations well[13]. Dr. Andrey Vyshedskiy has provided the group with a library of young child (2-5 years)[14] and the AI sub-group will develop

a more efficient algorithm to process the sound utilizing tensor flow. Eugene is considering looking at vowel sounds instead of phonemes which is the traditional linguistics approach to analyzing vocalizations, and Maurie looked into how the group can utilize the concept of vowel – formants to analyze vowel sounds more effectively. Formants are specific bands of frequency that determine the phonetic quality of a vowel. F1 is the first harmonic/fundamental. F2 is the second harmonic/fundamental. From the DSP or spectrogram data, F1 and F2 can be determined and together the two formants describe the vowel sound the person is making, i.e. there is a unique range of F1 and F2 values for each distinct vowel sound. So in the end, when the robot-software integration sub-group has finished their work in securing maintainable connection between the android and the robot devices, the team can implement the best sorting algorithm. Previous attempts with the project were unsuccessful partially because the connection between devices was not stable, the main goal of the robot-software integration sub-group is to secure connection between devices.

Child Vocalizations are imprecise.: As explored earlier, child vocalizations are imprecise. Also no one (adults and people without ASD) pronounces the same two phonemes or words exactly the same time after time, reasons being that pronunciation of any single sound is influenced by the neighboring sounds, so variation depends on the sentence being produced or the neighboring phonemes[15]. With regards to the program, the program needs to be able to address these ranges in a meaningful manner.

Milestone 2: "1. A successful algorithm will detect 95% of children's vocalizations that are identified as words by a human listener. 2. Furthermore, a successful algorithm will demonstrate higher correlation (better than 0.8) between the automatically calculated percent similarity and the subjective similarity score specified by the adults."[16]

Specific Aim 3: Provide a language therapy curriculum.
The programmed robot package will provide an expressive language curriculum and allow for dynamic interaction, allowing meaningful, interactive and self-paced practice outside of traditional clinical therapy sessions. With 200 words and phrases in the database, the robot will start on a low level and help the child practice with simple words. Eg. Mom, dad, dog, cat, etc. The curriculum will be cumulative and so the child will be quizzed with concepts from all past levels to cement understanding.

Aim 3.1. Create an Android application to control the robot via voice commands. [17]
The robot will be controlled by the android application, so the android application needs to both be able to understand and process the child's vocalizations and be able to connect with the robot device. The robot-software-integration team will both integrate the AI software and develop a secure connection via Blue-Tooth technologies. i.e. the integration group needs to guarantee a secure pathway for the devices to effectively "talk" to each other.

References

[1] Vyshedskiy, Andrey. *Early Language Assistance Therapy (ELAT) Robot for Children with ASD*. Pg.1-2
[2] Build Apps." *WowWee Labs - Latest SDKs for WowWee Products*, wowwee.com/about/labs
[3] Center for Disease Control. "Autism Spectrum Disorder (ASD)." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 26 Apr. 2018, www.cdc.gov/ncbddd/autism/data.html.
[4] Center for Disease Control. "Autism Spectrum Disorder (ASD)." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 26 Apr. 2018, www.cdc.gov/ncbddd/autism/data.html.
[5] Golinkoff, Roberta Michnick, et al. "Language Matters: Denying the Existence of the 30-Million-Word Gap Has Serious Consequences." *Child Development*, Wiley/Blackwell (10.1111), 13 Aug. 2018, onlinelibrary.wiley.com/doi/abs/10.1111/cdev.13128.
[6] Jones, Debra M, et al. "Rural and Remote Speech-Language Pathology Service Inequities: An Australian Human Rights Dilemma." *Taylor & Francis*, International Journal of Speech and Language Pathology, www.tandfonline.com/doi/full/10.1080/17549507.2018.1400103.
[7] ASHA. "Responding to Changing Needs in the 21st Century." *Averican Speech-Language-Hearing Association*, ASHA, www.asha.org/academic/reports/changing/.
[8] https://www.ulm.edu/slp/documents/supply-demand-slp.pdf
[9] "TensorFlow." *Wikipedia*, Wikimedia Foundation, 10 Oct. 2018, en.wikipedia.org/wiki/TensorFlow.
[10] *YouTube*, Www.tensorflow.org, 15 Feb. 2017, youtu.be/mWl45NkFBOc.
[11] University of Manitoba. "Formants." *Acoustic Phonetics: Formants*, University of Manitoba, home.cc.umanitoba.ca/~krussll/phonetics/acoustic/formants.html.
[12] "Build Apps." *WowWee Labs - Latest SDKs for WowWee Products*, wowwee.com/about/labs.
[13] Vyshedskiy, Andrey. *Early Language Assistance Therapy (ELAT) Robot for Children with ASD*.a pg. 4
[14] Vyshedskiy, Andrey. *Early Language Assistance Therapy (ELAT) Robot for Children with ASD*. Pg. 4
[15] https://www.cs.bham.ac.uk/~pxc/nlp/NLPA-Phon1.pdf
[16] Vyshedskiy, Andrey. *Early Language Assistance Therapy (ELAT) Robot for Children with ASD*. Pg. 4
[17] [17] Vyshedskiy, Andrey. *Early Language Assistance Therapy (ELAT) Robot for Children with ASD*. Pg.4