Personal Robots Group
MIT Media Lab

Speech Processing and Auditory Perception Lab
UCLA

Signal, Speech, & Language Interpretation Lab
University of Washington

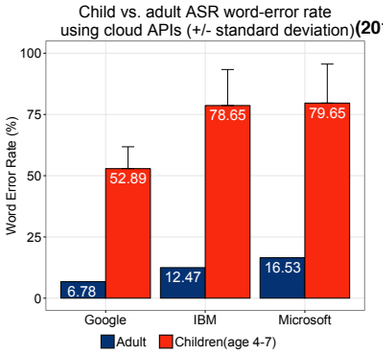
Development, Deployment and Evaluation of Personalized Learning Companion Robots for Early Literacy and Language Learning

HaeWon Park, Samuel Spaulding, Huili Chen, Cynthia Breazeal, MIT
Gary Yeung, Amber Afshan, Abeer Alwan,
Marlen Quintero, Alejandra Martin, Alison Bailey, UCLA
Mari Ostendorf, UW

Abstract

Social robot learning companions hold great promise for augmenting parents and teachers to promote childhood learning by physically, socially, and emotionally engaging with children. One of the most important factors for language skill development is sufficient exposure to a rich variety of spoken language and vocabulary – critical precursors to learning to read. The social context of exposure is also critical to concept development and the learning experience, i.e., simply hearing language is not enough, children need to actively participate and be emotionally and physically engaged to maximize their learning gains. Through this project, we are developing **a fully autonomous, collaborative, peer-like social robot system** with effective educational activities building on top of: **personalization algorithms** for story customization and dialogic question generation, **multi-modal assessment algorithms**, and, crucially, **Automatic Speech Recognition (ASR) and Spoken Language Understanding (SLU) systems for young children's speech**

"[Child-robot] Interactions relying solely on [Speech Recognition are] still out of reach"[1][2]



- **Years 1-2:** Gather corpus of child speech from target domains.
- **Years 2-3:** Develop robust tools for analyzing and modeling children's speech.
- **Years 3-4:** Conduct in-school and in-home deployments of personalized, interactive robot tutors that engage children in speech-based practice to promote literacy and language skills

[1] J. Kennedy et al. (2017). **Child Speech Recognition in Human-Robot Interaction: Evaluations and Recommendations**. Proc. of HRI 2017

[2] G. Yeung and A. Alwan, "**On the difficulties of automatic speech recognition for kindergarten-aged children**," in Proc. of INTERSPEECH 2018

Tasks and Objectives

GFTA-3: A Robot-Administered Articulation Test for Autonomous Child Speech Collection

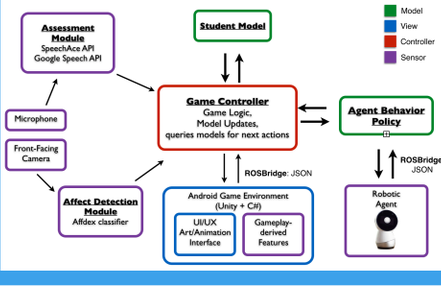
Digital systems to administer and score children's language comprehension and production have several advantages:
-Efficiency saves teachers' and students' valuable instructional time.
-Uniformity can mitigate discrepancies and biases in administration and interpretation of results.
-Scalability allows for the collection of much larger data sets from a greater diversity of speakers

- **Developed** a large collection of multimedia assets, including graphics, expressive recorded speech, and robotic animations, to support a child-robot speech-collection protocol.
- **Implemented** and **piloted** at 3 school sites in both Los Angeles and Boston, collecting over a dozen hours of child's speech recording.
- **Validated** that pre-K and kindergarten students were generally willing to accept the robot as an interface to complete oral language and speech assessments.
- **Identified** improvements in system and interaction design to elicit better speech samples.

"**Towards the development of personalized learning companion robots for early speech and language assessment**"
Gary Yeung, Amber Afshan, Marlen Quintero, Alejandra Martin, Samuel Spaulding, Hae Won Park, Alison Bailey, Cynthia Breazeal.
(Under submission)

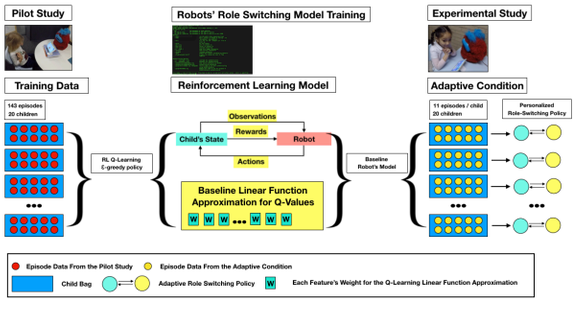
A cloud-based system architecture for long-term school and home-based deployments of social robot companions

Developing a common platform for autonomously analyzing and assessing children's speech and pronunciation during interactive word games between social robots and children.



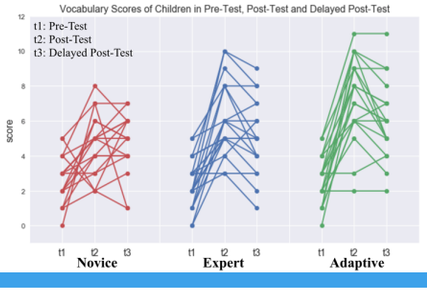
WordQuest

A collaborative vocabulary game with a robot using an Active Role Adaptation model



Robot tutors often take a fixed **role** in an interaction, as either a teacher-like **expert** or a peer-like **novice**. We developed an **RL-based role-switching policy** that adapts the robots role and behaviors each turn, based on real-time interaction data.

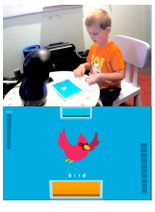
Over two learning sessions, children played with either a **fixed-Expert**, **fixed-Novice**, or **Adaptive Role** robot. We found that children in the adaptive condition had increased scores on both an immediate and delayed post-experiment vocabulary test

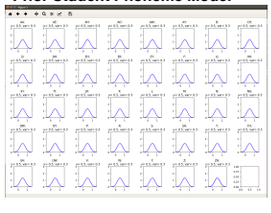


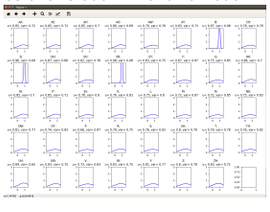
Active Role Adaptation for Social Robot Learning Companion.
Chen, H., Park, H.W., Dong, X., Breazeal, C. (2019), (Under Submission)

WordRacer

A competitive speech game against an agent that models pronunciation via Gaussian Processes + Active Learning

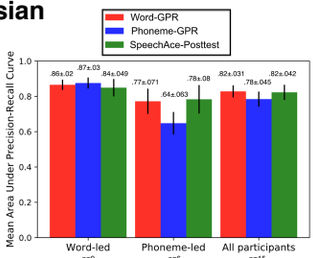






We developed an autonomous robot that **models children's word pronunciation via Gaussian Process Regression (GPR), augmented with an Active Learning protocol** that determines the game content.

We show that the system is capable of supporting a **speech-based game** interaction while **accurately assessing children's pronunciation** ability, with ground truth determined by a post-experiment evaluation by human raters.



"A Social Robot System for Modeling Children's Word Pronunciation."
Spaulding, S., Chen, H., Ali, S., Breazeal, C Proc. of AAMAS 2018.

This research was supported by the National Science Foundation (NSF) under Grant IIS-1734443.