# Development, Deployment and Evaluation of Personalized Learning Companion Robots for Early Literacy and Language Learning (NSF 1734380)

MIT: Cynthia Breazeal, Hae Won Park
UCLA: Abeer Alwan, Alison Bailey
UW: Mari Ostendorf

https://robot-literacy-tools.media.mit.edu/

This NRI project will develop, deploy and evaluate **personalized companion robots** to assist kindergarten-age children in learning language and vocabulary skills. The aim is to accelerate the impact of social robots for **early childhood education** in schools and at home. The project will generate new insights for how to develop expressive, socially responsive robots that provide more effective, engaging, and empathetic educational experiences for young children.

## Robot Cognitive Architecture

### Long-term Personalized Reading Companion

#### Key Problem and Significance

**Automatic Reading Skill Assessment and Learning Content Personalization**
- Personalized intervention is known to be the most effective method for early literacy and language learning but it is hard to achieve in classroom setting

**Personalizing Robot Behavior to Maximize Engagement**
- Each student learns and is motivated differently, so the robot learning companion should personalize its behavior policy to maximize each student's engagement.

**Reliable, Robust, Affordable Long-term Robot Platform**
- Demonstrate and evaluate a long-term deployable robot system in-school and at-home.

#### Scientific Impact

- Personalizing learning content has been explored by many prior works, however, personalizing the learning experience by monitoring children's engagement through a peer-like learning companion has high impact in education research, learning science, and human-robot interaction.
- A reliable, robust, affordable long-term deployable robot system is in high demand for our research community, and we plan on open sourcing the architecture.
- With the help of long-term deployable robot platform, we are able to conduct true real-world user research to measure long-term effect of robot companions.

#### Solutions (Approaches, innovations, Contributions)

***Modeling Children's Word Knowledge & Pronunciation***

Approach:
- An interactive tablet game (WordRacer) that facilitates child-robot word pronunciation game play
- Word- and phoneme-level student models based on Gaussian Process Regression (GPR)
- An active learning protocol for efficiently determining and introducing adaptive, personalized content

Innovations: A novel attempt to rigorously and autonomously model children's spoken word pronunciation in the context of an interactive game with a social robot [4,5].

***Personalized RL Robot Peer Behavior Policy***

Approach:
- An interactive tablet game (WordQuest) that facilitates child-robot peer-to-peer word learning
- Models for the robot's different collaborative roles (e.g., tutor, tutee) and a set of behaviors associated with each role (e.g., question asking, encouragement, providing information, etc.)
- Robot learns RL based personalized role-switching behavior policy that maximizes each child's learning performance

Innovations: Novel RL based robot adaptive role-switching behavior policy [6,7].

Contributions: Multimodal data collection protocol and infrastructure for capturing children's pronunciation, facial expression, touch actions on the tablet (e.g., dragging, panning), in-game performance (e.g., successful attempts, success rate), and the robot's verbal and nonverbal behaviors (e.g., question asking, demonstration, etc.). Evaluation of learning companion interactions.

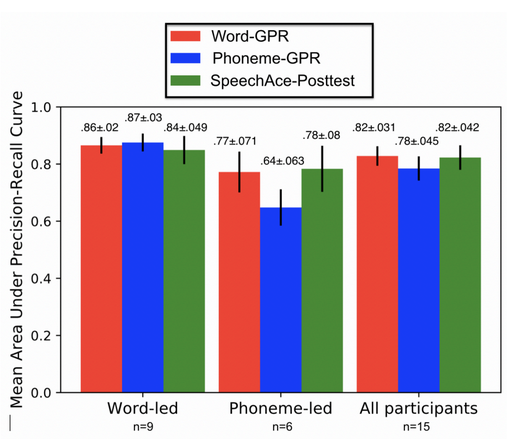


**Figure.** Classification results (total area under the Precision-Recall curve) of Word- and Phoneme-GPR models trained on data from the Experiment phase. SpeechAce results were derived directly from Posttest data.

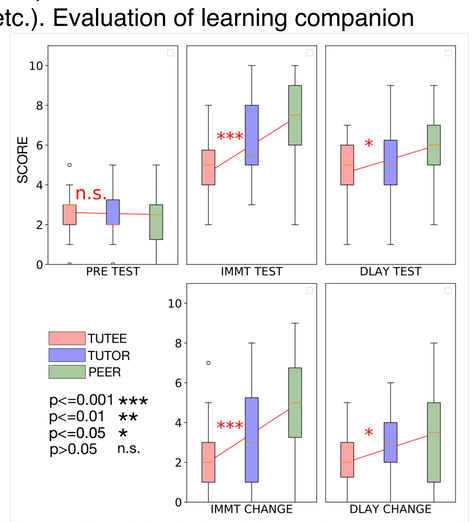


**Figure.** Trend Analysis of the word test scores of children's vocabulary assessment. A trend shows the significant advantage of the adaptive role-switching policy (PEER condition).

*Question Proposal*

*Verbal & Non-verbal*

*Training Data*

### Dialogic QnA

#### Key Problem and Significance

**Accounting for child engagement and cognitive state in the interaction strategy**
- Understanding the child (not just the words) is important for appropriate timing and intent of robot responses

**Dynamic question generation based on story context & child interest**
- Children differ in what captures their attention, so the system needs to be capable of generating a question at any point in the story

#### Scientific Impact

- Dialog systems can benefit from understanding users of any age, and core principles and computational frameworks should generalize to other scenarios
- Contextually grounded question generation frameworks are relevant to robot learning as well as child learning

#### Solutions (Approaches, innovations, Contributions)

***Understanding child engagement & cognitive state***

Approach: Multimodal QnA Timing
- Leverage visual cues (engagement) and prosodic & disfluency cues (cognitive state)
- Integrate cues to identify timing of robot responses (turn-taking) & response intent (encourage, hint, factual question, open question, etc.)

Innovations:
- Use of prosodic cues to detect child uncertainty as a point for encouragement
- Analysis of child disfluencies as cues to improve turn-taking and intent selection

Contributions: New corpus of disfluency annotated children's speech

***Contextually grounded question generation***

Approach:
- Represent context using story graphs and entity cache
- Automatically generate questions and answers for multiple points in the story and multiple entities
- Using story context as grounding, compare sequence-to-sequence vs. dependency-rewrite approaches for question/answer generation

Innovations: Novel neural network architecture for question-answer generation

Contributions: New corpus of story questions designed for spoken conversations

*txt/audio Time Alignment*

### Automatic Child Speech Recognition

#### Key Problem and Significance

**Development of an Engaging Speech Collection Protocol Administered by a Social Robot**
- The robotic system is capable of administering the GFTA3 test, a popular clinical speech assessment protocol as well as other tasks for reading and listening comprehension

**Analyses and development of effective child ASR**

#### Scientific Impact

Children's speech is more variable than adult speech. There is a lack of DB that captures the variability to study how children's speech changes with age. Our DB and analyses shed light on this important issue [1-3].

*Real-time ASR*

*Training Data*

#### Solutions (Approaches, innovations, Contributions)

Approach: Extensive data collections with the goal of obtaining at least 200 hours of child (ages 4-8) speech

Innovations: The data is collected longitudinally, that is, a subset of the children will be recorded every year so that we can observe and analyze how their speech changes as they grow. This would be the first database of its kind to offer longitudinal information.

Contributions:
- New corpus of child-robot speech with a subset being monitored longitudinally
- New Automatic Speech Recognition algorithms that are robust to the variability evident in child speech

## Broader Impact

- **Engaging Teachers and Parents: Outreach, Guidelines, and Best Practices** - 1) Focus group sessions in which all stakeholders discuss experiences with implementation of the social robots in order to both improve the robotic functionality, as well as determine under what circumstances implementation is most effective for students.
  2) Review of third-party video of assessment sessions during which student level of engagement and language and literacy performance in different settings and with different partners can be evaluated. From these efforts we anticipate isolating best practices and establishing guidelines that we will freely distribute.
- **Datasets, Tools, Education App**s: a set of databases, long-term robot platform tools, and educational activities developed for use with the social robot platform will be disseminated to the research and education community.
- **Cross-Disciplinary Training for Students**: cross-disciplinary training through Annual Workshops including topics: reliability and validity of technology-based systems, child development and early education, performance assessments, advanced user interfaces and technologies, experimental design and analysis, spoken language technologies, etc.

[1] G. Yeung, A. Afshan, M. Quintero, A. Martin, S. Spaulding, H. W. Park, A. Bailey, C. Breazeal, and A. Alwan, "Towards the development of personalized learning companion robots for early speech and language assessment," AERA 2019.
[2] G. Yeung and A. Alwan, "On the difficulties of automatic speech recognition for kindergarten-aged children." in Interspeech, 2018.
[3] G. Yeung, S. M. Lulich, J. Guo, M. S. Sommers, and A. Alwan, "Subglottal resonances of American english speaking children," Journal of the Acoustical Society of America, 2018.
[4] S. Spaulding, H. Chen, S. Ali, M. Kulinski, and C. Breazeal, "A social robot system for modeling children's word pronunciation: Socially interactive agents track," AAMAS 2018.
[5] I. Grover, H. W. Park, and C. Breazeal, "A semantics-based model for predicting children's vocabulary," in IJCAI 2019.
[6] H. Chen, H. W. Park, and C. Breazeal, "Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement," Computers & Education (to appear), 2020.
[7] H. Chen, H. W. Park, X. Zhang, and C. Breazeal, "Impact of interaction context on the student affect learning relationship in child-robot interaction," in HRI 2020 (to appear).