Can a Signing Virtual Human Engage a Baby's Attention?

Setareh Nasihati Gilani, David Traum
Rachel Sortino
Grady Gallagher
Kailyn Aaron-Lozano
Cryss Padilla, Ari Shapiro, Jason Lambertson
Laura-Ann Petitto, PI*

Funding: National Science Foundation (NSF) INSPIRE IIS-1547178 (L-A Petitto, PI), W.M. Keck Foundation (L-A Petitto PI), NSF SBE-1041725 Science of Learning Center, Visual Learning & Visual Learning (L-A Petitto, Co-PI)

A Crisis in Early Development

- Age 6-12 months is a **critical period** for language development
- Many babies are at risk for **minimal language exposure**

Goal: To facilitate language learning in young babies

Why can't babies learn language from a TV?

Attention

Research has assumed babies do not learn language from a screen* because they need longer general attention

We hypothesized babies need *specific* language patterns for that time of brain development

Social Interaction

Research has assumed that environmental social interactions are vital to language acquisition in babies

We hypothesized babies need language samples that are Socially Contingent

*Krcmar, 2011; Krcmar, et. al. 2007; Richert et. al. 2011

A change in terminology

Engagement

System needs the natural phonetic patterns found in all human language

Social Contingency

System needs to receive input from babies about their level of engagement

Innovative Solutions

Engagement

Avatar produces language constructed with the brain's peaked sensitivity to the rhythms of language within ages 6-12 months

Social Contingency

Use thermal imaging to identify babies' engagement = *internal* readiness to learn, triggering socially contingent dialogues

Our Hypotheses

Avatar's Natural Language Patterns

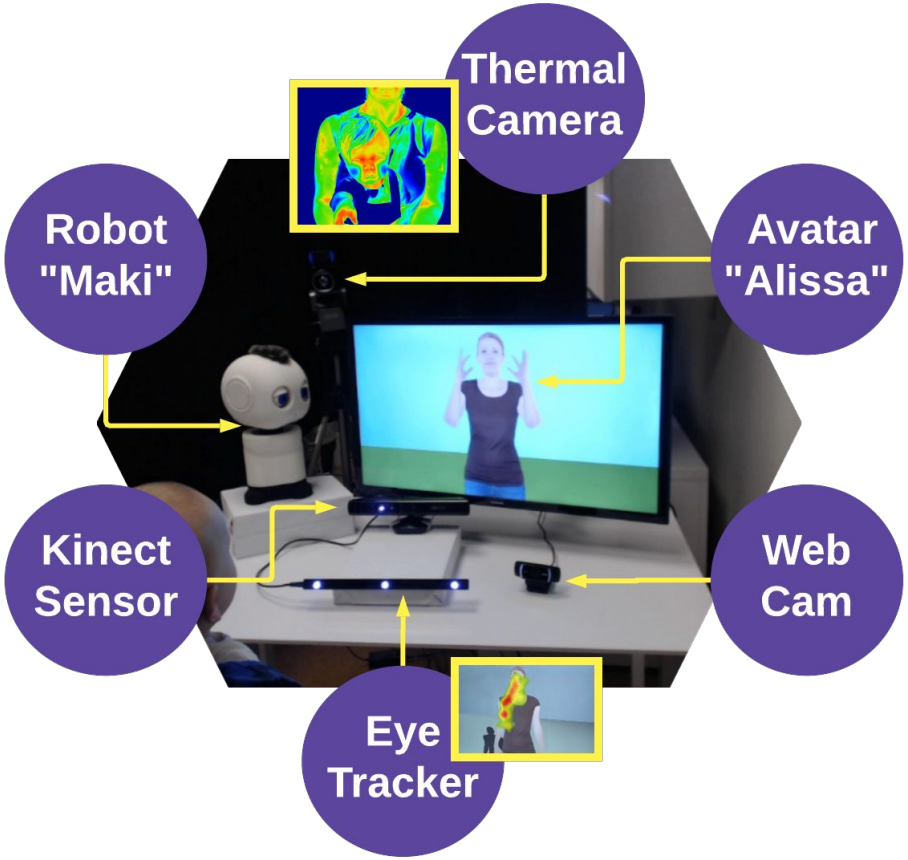
+

Thermal + socially contingent dialogues

= Language Learning

Introducing, RAVE

Robot Avatar thermal Enhanced language learning tool prototype



What does the Avatar do?

Nursery Rhymes

Socially Contingent Responses

Experimental Questions

1. Do **babies attend** to the avatar and **respond** to its communicative behaviors?
2. Can **babies recognize** different avatar behaviors?
3. Can an **avatar elicit** language responses from babies?
4. Do the **parents impact** the conversational interactions?

Participants

- T=250; 4 babies
- Ages 7 - 10 months
- One baby: Sign-exposed (American Sign Language)
- Three babies: Non-sign-exposed

Methods

Familiarization

Condition 1

Condition 2

Babies' Spontaneous Types of Responses

1. Sustained Visual Attention
2. Linguistic Responses
3. Social/Gestural Responses

Results: Overall Response Rates

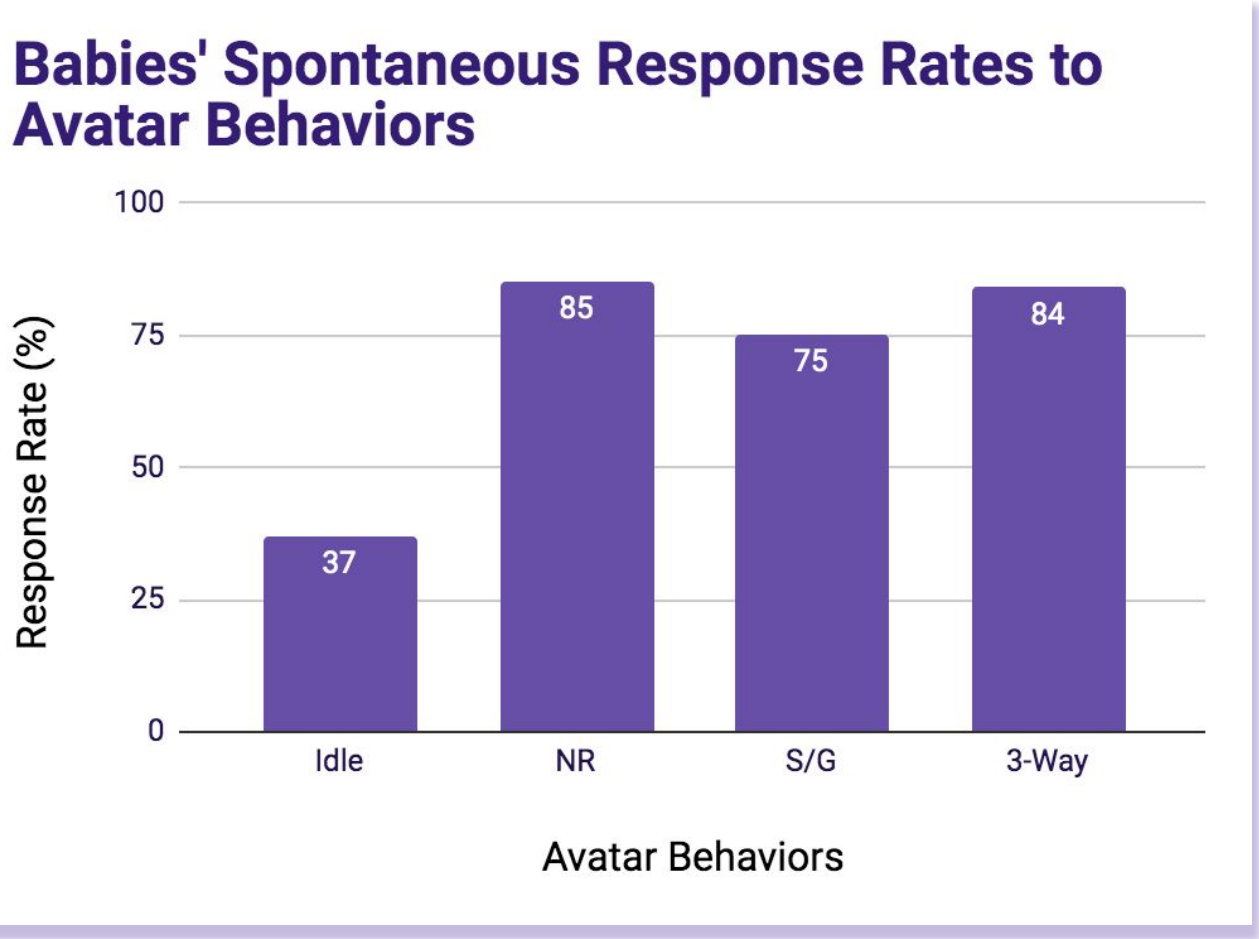
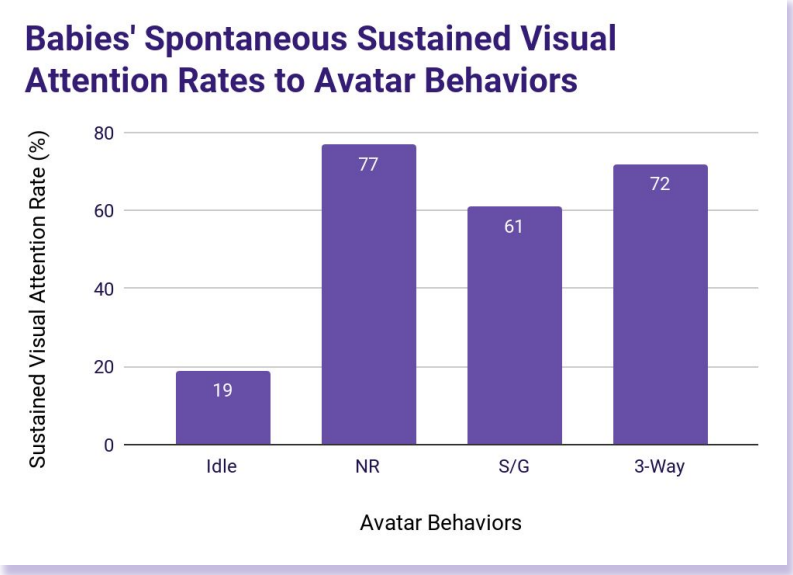


Figure 1 Babies produced significantly more behaviors towards Avatar's actions (NR, S/G, 3-way) than to its idle ($t=3.35$, $p=0.01$)

Results: Sustained Visual Attention



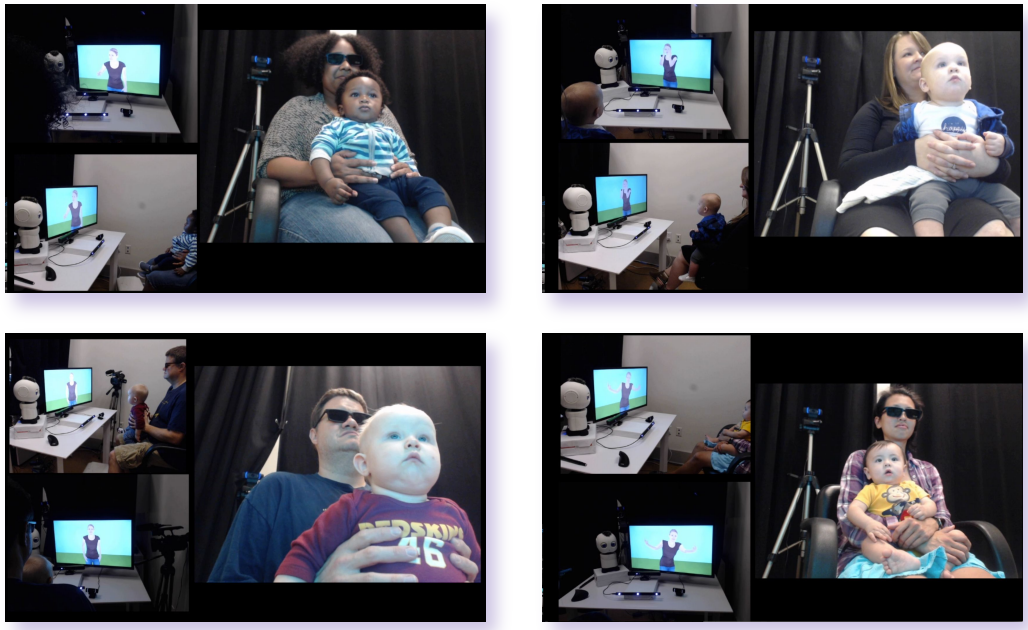


Figure 2 Babies produced sustained visual attention most to the avatar's linguistic nursery rhymes and least to its idles

Results: Linguistic Response

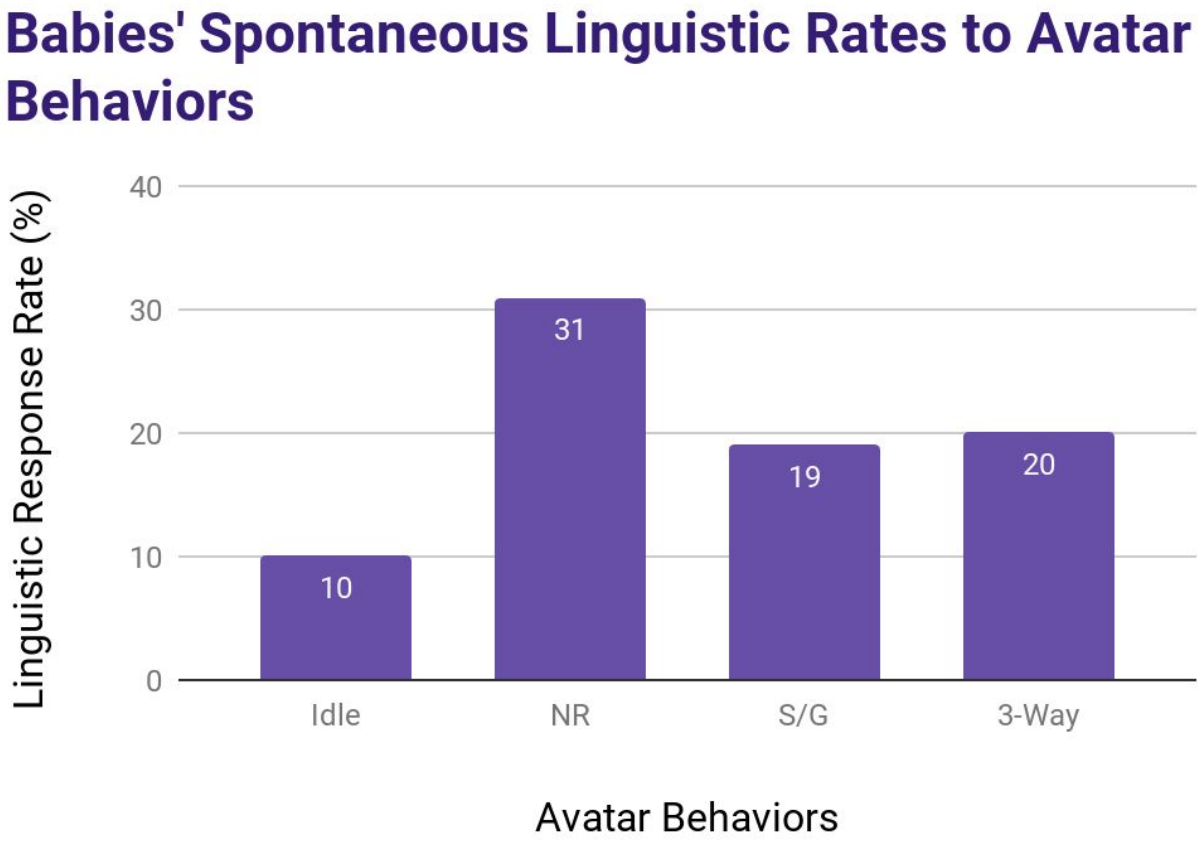


Figure 3 Babies produced more linguistic responses to Avatar's Nursery Rhymes than its other actions and idle

Results: Parent Intervention

From *absence* (condition 1) to *presence* (condition 2) of parental involvement

- Overall responses **decreased** significantly ($t=2.22$, $p < 0.05$)
- SVA **decreased** significantly ($t=4.3$, $p < 0.005$)
- Linguistic responses **increased** significantly ($t=2.4$, $p < 0.05$)

Discussion

1. Babies **attention** evidenced a preference for the avatar's linguistic and communicative behaviors over non-communicative idles
2. Babies showed evidence that they **recognized** different linguistic, communicative, and non-communicative avatar behaviors
3. The avatar's Linguistic Nursery Rhymes **elicited** the highest rates of linguistic productions
4. Parent interference **decreased** sustained visual attention but **increased** linguistic responses

Conclusions

- Most babies had **no prior exposure to sign language**, yet they were captivated by the avatar and exhibited spontaneous, differential engagement with the avatar
- This work provides a novel demonstration of the potential for avatars to facilitate language learning in **young deaf babies**

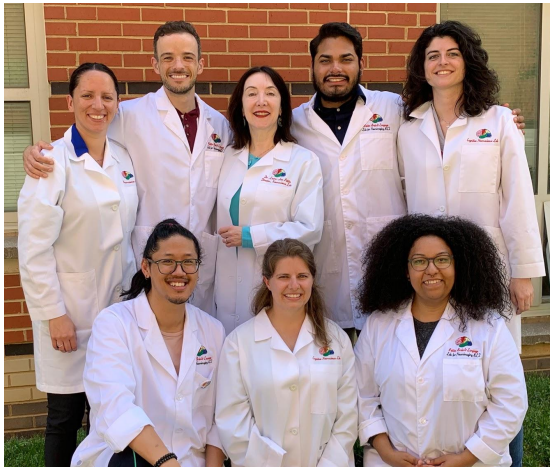
Beyond sound, all babies are sensitive to the rhythmic patterns of language - showing that Deaf babies must have early exposure to signed languages within critical period as a bridge to learning multiple languages




Merci Beaucoup! Thank you!

Funding: National Science Foundation (NSF) INSPIRE IIS-1547178 (L-A Petitto, PI), W.M. Keck Foundation (L-A Petitto PI), NSF SBE-1041725 Science of Learning Center, Visual Learning & Visual Learning (L-A Petitto, Co-PI)

PI Science Director: Dr. Laura-Ann Petitto



Our BL2 team at Gallaudet University

References: See Petitto.net

End

Mio Amico - next generation

Innovations !

Needed a way

- to grip ATTENTION = Hypothesis make Avatar produce language constructed with the brain's peaked sensitivity to the rhythms of language within ages 6-12 months
- to interact in meaningful ways (based on baby's internal states) = Hypothesis make Avatar produce language that is based on babies' emotional engagement, thereby simulating a conversation that is meaningful & contingent on the babies' internal states
- Innovations
    - Nursery Rhymes with timing of brain's sensitivity to language phonology
    - Thermal IR imaging: other AI systems only have a way to "Read" a baby superficially through external measures of Eye Gaze. By contrast we are reading internal states (with thermal)
    - Thermal Triggers Avatar Stop & Start of conversation integrated with
    - Social Dialogue Scripts