A Comparison of Children Learning New Words from Robots, Tablets, & People

Jacqueline Kory Westlund, Leah Dickens, Sooyeon Jeong, Paul Harris, David DeSteno, & Cynthia Breazeal

[a]MIT Media Lab, [b]Northeastern University, [c]Harvard University

Abstract. This work investigates young children's perceptions of social robots in a learning context. Because social robots are a relatively new technology, direct comparison to more familiar means of learning could give us useful insights. Here, we compared the efficacy of three sources of information (human, robot, and tablet/iPad) with respect to children's rapid learning of new words. Our results suggested that in this simple case, all three interlocutors served equally well as providers of new words. However, children strongly preferred learning with the robot, and considered it to be more like a person than like an iPad. Follow-up work will examine more complex learning tasks.

Keywords. Education; learning; children; social robots.

INTRODUCTION

The development of children's early oral language skills is critical for nearly all subsequent learning. Differences in children's early vocabulary ability can predict differences in reading ability in middle and high school [1], which could magnify over time, inhibiting later growth [2]. Given the importance of language, it would be beneficial to find new ways to supplement the education of children who may not currently be getting enough support, instruction, or practice. We suggest that emerging technologies can help fill this gap.

Computers, tablets, iPads, and even robots are being introduced in many educational settings [3]. Technology has the advantages of being easily customizable, adaptive to individual learners, as well as broadly deployable. But despite the frequent success of these technologies, we often intuitively assume that humans have some "special sauce" that makes us more suited to being teachers and learning companions than any kind of technology. This may be especially true with regards to learning language, which, as a socially situated medium that is *for* sharing meaning, still seems a uniquely "human" ability.

To this end, we are exploring the effectiveness of technology, specifically robots, as language learning companions for children. Robots occupy a unique role because their embodiment allows them to employ more of the "human" behaviors and social cues that are recognized as crucial in language learning [4]. Children seem to readily learn words from both mobile devices [5] and robots [6], [7]. However, one concern about some of these prior studies is that the learning conditions presented may not reflect children's usual language learning, which often proceeds rapidly and without feedback from a teacher. As such, in this work, we focus on one particular type of rapid, albeit approximate, word learning without feedback, known as "fast mapping" [8]. Although grasping the full meaning of a new word can take time, the initial mapping is often accomplished quickly. Accordingly, we ask whether children display a process of fast mapping with a social robot or a tablet, just as they would with a human interlocutor. We expected that children would learn equally well from the human and robot, and that the tablet would fair somewhat worse due to its lack of social embodiment. Furthermore, we probed children's perceptions of the robot in an attempt to understand how they construed it. The study is modeled closely on the procedure in [9].

METHODS

Nineteen children ages 4-6 (10 female, 9 male), from a Greater Boston area preschool serving a mainly middle-class population participated in two sessions, set about one week apart. The experiment followed a within-subjects design.

In Session 1, children were first asked questions about whether they thought a robot was more like a person or like an iPad. Then, each child looked at three series of ten pictures of unfamiliar animals, presented one image at a time on the tablet. They viewed ten pictures with just the tablet, ten with the robot (Figure 1), and ten with the second experimenter (thirty total). The order of the interlocutors was counterbalanced to handle order effects. The order in which the pictures were presented was held constant across interlocutors. A Samsung Galaxy Tablet was used to present the animal pictures. When the tablet was the interlocutor, recorded human speech was played back through the tablet's speakers. The robot was a DragonBot [10], which was teleoperated by a second experimenter.

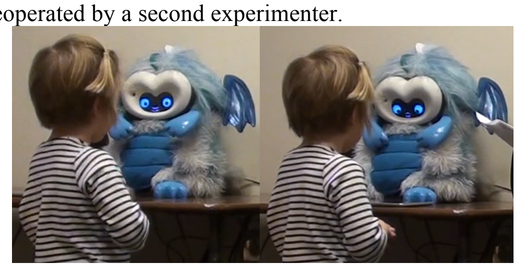

Figure 1: Children viewed pictures of novel animals with the DragonBot as well as with a person or with the tablet.

During the picture viewing, the child's interlocutor commented positively but uninformatively on the animal shown for 8 of the 10 pictures, e.g., "Look at that!" The remaining two animals were named, e.g., "Ooh, a kinkajou! See the kinkajou?" This presented the opportunity for fast mapping to occur. After each set of pictures, we measured children's learning with a recall test. Finally, we asked the earlier questions again, and probed children's preferences for learning from the human vs. robot vs. iPad.

In Session 2, we wanted to see whether children's thoughts about robots had changed, and to test retention of the animal names they had learned. They were given the same recall tests and were asked the same sets of questions.

RESULTS

We found that, across the three conditions, children learned a mean of 4.3 of the 6 animals correctly (71.7%

correct, *SD*=1.84). However, there were no significant differences across conditions in how many names were learned. In Session 2, children's retention was nearly as good, naming a mean of 3.9 of 6 animals correctly (65.0% correct, *SD*=1.48), indicating that they did learn the names.

Children expressed a strong preference for learning with the robot. After Session 1, 63.2% (12 of 17) children preferred the robot, 1 child preferred the iPad, 1 preferred the person, and 5 liked all three equally (two children were not asked this question in Session 1). After Session 2, 73.7% (14 of 19) children preferred the robot; 2 preferred the iPad, and 3 liked all three equally. Thus, although learning success appeared the same, enthusiasm was higher for the robot.

Regarding children's perceptions of the robot, the most telling questions were *"When a robot answers a question, is it more like a person or more like an iPad?"* and *"When a robot teaches you something…"* Prior to interacting with the robot, children were split in their answers ("Answers…": 52.6% person, 47.4% iPad; "Teaches…": 47.4% person, 52.6% iPad). After interacting, more children thought the robot was more like a person ("Answers…": 78.9% person, 21.1% iPad; "Teaches…": 68.4% person, 31.6% iPad). However, during the follow-up Session 2, some children reverted back to their original opinion ("Answers…": 36.7% person, 63.2% iPad; "Teaches…": 68.4% person, 31.6% iPad). For the remaining questions, children generally thought the robot was more like a person.

DISCUSSION

We examined the efficacy of, as well as children's subjective attitudes toward, three different sources of information (human, robot, and tablet) with respect to word learning. Our results suggested that in this simple case, contrary to our hypotheses, all three interlocutors served equally well as providers of novel animal names. We suspect that this is due to the simple nature of the learning task. When only one picture is shown and named, children need not observe the interlocutor's social cues to understand what is being referred to by the novel name that is provided. Given that the key benefit provided by the robot and human over the tablet is their ability to offer social cues, it is understandable that, because these cues were not necessary, the tablet was equally well suited to the learning task.

However, children showed a clear preference for learning with the robot. Their enthusiasm and, therefore, likely engagement was higher with the robot. It is unclear whether this was merely a novelty effect. We suspect that given a sufficiently interesting activity with the robot, children's preference for a robot over a tablet would not simply be novelty – recent work has shown that children can remain interested and engaged with a robot during educational games for a month or more [6], [7].

Regarding children's perception of the robot, our results suggest that although children initially expect a robot to engage them just like any other technological tool, their perceptions of it rapidly change. Note that this shift was evident for the two questions in which children were invited to appraise the robot as an active, social partner, i.e., as an interlocutor that is able to teach and answer questions. They come to perceive it as being more "human," more like a *someone* than a *something,* which suggests that they will attend to its social cues when they need to learn.

Follow-up work is now in progress to probe the social dimension farther. We are looking at tasks that require social information for learning (e.g., gaze) and more closely mirror what happens in "real-life", such as when a child needs to determine which of multiple target objects is the referent. Because robots can operate in the same spaces that we do (while tablets are limited to a two-dimensional screen world), it is an interesting challenge to identify clear differences between the social capabilities of a human and a robot. Our future work will continue exploring how children learn from different agents, and which social cues are truly important for learning.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (NSF) under Grant 122886 and Graduate Research Fellowship Grant No. 1122374. Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not represent the views of the NSF.

REFERENCES

1. M. Fish and B. Pinkerman, "Language skills in low-SES rural Appalachian children: Normative development and individual differences, infancy to preschool," *J. Appl. Dev. Psychol.*, vol. 23, no. 5, pp. 539–565, 2003.
2. B. Hart and T. R. Risley, *Meaningful differences in the everyday experience of young American children.* ERIC, 1995.
3. C. K. Blackwell, A. R. Lauricella, E. Wartella, M. Robb, and R. Schomburg, "Adoption and use of technology in early education: The interplay of extrinsic barriers and teacher attitudes," *Comput. Educ.*, vol. 69, pp. 310–319, Nov. 2013.
4. P. L. Harris, "Trust," *Dev. Sci.*, vol. 10, pp. 135–138, 2007.
5. S. Judge, K. Floyd, and T. Jeffs, "Using Mobile Media Devices and Apps to Promote Young Children's Learning," in *Young Children and Families in the Information Age*, K. L. Heider and M. R. Jalongo, Eds. Springer Netherlands, 2015, pp. 117–131.
6. J. Kory Westlund and C. Breazeal, "The interplay of robot language level with children's language learning during storytelling," in *Proceedings of the 2015 ACM/IEEE international conference on Human-robot interaction*, Portland, OR, 2015.
7. J. Movellan, M. Eckhardt, M. Virnes, and A. Rodriguez, "Sociable robot improves toddler vocabulary skills," in *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, 2009, pp. 307–308.
8. S. Carey, "Beyond Fast Mapping," *Lang. Learn. Dev. Off. J. Soc. Lang. Dev.*, vol. 6, no. 3, pp. 184–205, Jul. 2010.
9. L. Markson and P. Bloom, "Evidence against a dedicated system for word learning in children," *Nature*, vol. 385, no. 6619, pp. 813–815, Feb. 1997.
10. J. M. Kory, S. Jeong, and C. L. Breazeal, "Robotic learning companions for early language development," in *Proceedings of the 15th ACM on International conference on multimodal interaction*, New York, NY: ACM, 2013, pp. 71–72.