



Child-Robot Interactions for Second Language Tutoring to Preschool Children

Paul Vogt*, Mirjam de Haas, Chiara de Jong, Peta Baxter and Emiel Krahmer

Tilburg Center for Cognition and Communication, Tilburg University, Tilburg, Netherlands

In this digital age social robots will increasingly be used for educational purposes, such as second language tutoring. In this perspective article, we propose a number of design features to develop a child-friendly social robot that can effectively support children in second language learning, and we discuss some technical challenges for developing these. The features we propose include choices to develop the robot such that it can act as a peer to motivate the child during second language learning and build trust at the same time, while still being more knowledgeable than the child and scaffolding that knowledge in adult-like manner. We also believe that the first impressions children have about robots are crucial for them to build trust and common ground, which would support child-robot interactions in the long term. We therefore propose a strategy to introduce the robot in a safe way to toddlers. Other features relate to the ability to adapt to individual children's language proficiency, respond contingently, both temporally and semantically, establish joint attention, use meaningful gestures, provide effective feedback and monitor children's learning progress. Technical challenges we observe include automatic speech recognition (ASR) for children, reliable object recognition to facilitate semantic contingency and establishing joint attention, and developing human-like gestures with a robot that does not have the same morphology humans have. We briefly discuss an experiment in which we investigate how children respond to different forms of feedback the robot can give.

Keywords: social robots, second language tutoring, education, child-robot interaction, robot assisted language learning

OPEN ACCESS

Edited by:

Mila Vulchanova,
Norwegian University of Science and
Technology, Norway

Reviewed by:

Ramesh Kumar Mishra,
University of Hyderabad, India
Vera Kempe,
Abertay University, UK

*Correspondence:

Paul Vogt
p.a.vogt@uvt.nl

Received: 26 October 2016

Accepted: 06 February 2017

Published: 02 March 2017

Citation:

Vogt P, de Haas M, de Jong C,
Baxter P and Krahmer E
(2017) Child-Robot Interactions for
Second Language Tutoring to
Preschool Children.
Front. Hum. Neurosci. 11:73.
doi: 10.3389/fnhum.2017.00073

SOCIAL ROBOTS FOR SECOND LANGUAGE TUTORING

Given the globalization of our society, it is becoming increasingly important for people to speak multiple languages. For instance, the ability to speak foreign languages fosters people's mobility and increases their chances for employment. Moreover, immigrants to a country need to learn the official host language. Since young children are most flexible at learning languages, starting second language (L2) learning in preschool would provide them a good opportunity to acquire the second language more fluently at a later age (Hoff, 2013).

One trend in the digital age of the 21st century is that technologies are being developed for educational purposes, including technologies to support L2 tutoring. There exist many forms of digital technologies for PCs, laptops or tablet computers that support second language learning, although there is little evidence about their efficacy (Golonka et al., 2014; Hsin et al., 2014). While children can benefit from playing with such technologies, these systems lack the situated and

embodied interactions that young children naturally engage in and learn from (Glenberg, 2010; Leyzberg et al., 2012). Social robots represent an emerging technology that provides situatedness and embodiment, and thus have potential benefits for educational purposes. In essence, social robots are autonomous physical agents, often with human-like features, that can interact socially with humans in a semi-natural way for prolonged periods of time (Dautenhahn, 2007). The use of social robots, in comparison to more traditional digital technologies, allows for the development of tutoring systems more akin to human tutors, especially with respect to the situated and embodied social interactions between child and robot. Thus, this offers the opportunity to design robots such that they interact in a way that optimizes the child's language learning.

Recently, an increasing interest has emerged to develop social robots to support children with learning a second language (Kanda et al., 2004; Belpaeme et al., 2015; Kennedy et al., 2016). While a social robot cannot provide tutoring to the level humans can, recent studies suggest that using social robots can result in an increased learning gain compared to digital learning environments for tablets or computers (Han et al., 2008; Leyzberg et al., 2012). It is, however, unclear why this is the case. Perhaps the physical presence of the robot draws the attention of children for longer periods of time, but the embodiment and situatedness of the learning environment perhaps also helps the children to ground the language more strongly than interactions with virtual objects do.

While there is a fair body of research on robot tutors, a comprehensive description of the design features for a second language robot tutor based on what is known about children's language acquisition is lacking. What are the design features of child-robot interactions that would support second language learning? And, to what extent can these interactions be implemented in today's social robot technologies? In this perspective article, we try to answer these questions based on theoretical accounts from the literature on children's language acquisition in combination with our own experiences in designing a tutor robot.

DESIGNING CHILD-ROBOT INTERACTIONS

In our project, we aim to design a digital learning environment in which preschool children interact one-on-one with a social robot that supports either their learning of English as a foreign language, or the school language for those children who have a different native language (Belpaeme et al., 2015). In particular, the project aims to develop a series of tutoring sessions revolving around three increasingly complex domains (numbers, spatial relations and mental vocabulary). In each session, the child will engage with the robot (a Softbank Robotics NAO robot) in a game-like scenario focusing on learning a small number of target words. The contextual setting is generally displayed on a tablet computer that occasionally also provides some verbal support, however, the robot acts as the interactive tutor. Below we discuss the design features and considerations that we believe are crucial to design a successful tutoring system.

Peer-Like Tutoring

One of the first questions that comes up when designing a robot tutor is whether the robot should take the role of a teacher or a peer. Research on children's language acquisition has demonstrated that children learn more effectively from an adult who can use well-defined pedagogical methods for teaching children using clear directions, explanations and positive feedback methods (Matthews et al., 2007). However, designing and framing the robot as an adult tutor has the disadvantage that children will form expectations about the robot's behavior and proficiency that cannot be met with current technology (Kennedy et al., 2015). Due to technological limitations of the robot and underlying software, communication breakdowns are more likely to occur than with a human. For a peer robot introduced as a fellow language learner, breakdowns in communication are more acceptable. Moreover, interacting with robots acting as peers is conceived as more fun (Kanda et al., 2004), allows for learning-by-teaching (Tanaka and Matsuzoe, 2012) and has a proven to be efficient in teaching children how to write (Hood et al., 2015). Furthermore, there is some evidence that children's learning can benefit from interacting with peers (Mashburn et al., 2009). Given these considerations, we believe it is desirable to frame or introduce the robot as a peer and friend, yet design its interactions insofar possible based on pedagogically well-established strategies to scaffold language learning.

First Impressions

To implement effective tutoring, the robot needs to interact with children in multiple sessions, so they have to be motivated to engage in long-term interactions with the robot. Establishing common ground between child and robot can contribute to this (Kanda et al., 2004), but first impressions to establish trust and rapport are also crucial (Hancock et al., 2011).

Despite the wealth of studies regarding the introduction of entertainment robots as toys to children (e.g., Lund, 2003), surprisingly little research has been conducted on designing protocols on how to introduce a robot tutor to a group of preschool children. Fridin (2014) presents one exception, and found that introducing a robot tutor to children in group sessions improved subsequent interactions compared to introducing the robot to children in individual sessions. Another study by Westlund et al. (2016) found that the way a robot is framed, either as a machine or a social entity, affected the way children later engaged with the robot. They concluded that introducing the robot as a machine could create a more distant relation between child and robot, thus reducing acceptance. We therefore decided to frame the robot in our project as a social playmate for the children and introduced the robot in a group session. However, the NAO robot is slightly taller and more rigid than the fluffy huggable Tega robot, which Westlund et al. (2016) used, and we observed that some 3-year-old children were somewhat intimidated by the NAO robot on their first encounter. Such a first impression of the robot could reduce the trust that the child had for the robot, which could negatively affect their willingness to interact with the robot in the short-term, but also in the long-term. To develop a successful first encounter and to build

trust between the child and robot, we designed the following strategy for introducing the robot to 3-year-old children at their preschool.

Pilot studies revealed that some children got anxious when the robot was introduced and then suddenly started to move. To familiarize children prior to their first encounter with the robot, it is therefore advisable to prepare them well. For our study, we sent coloring pages of the robot to the preschools during recruitment and asked the pedagogical assistants to talk a little bit about the robots to the children. About 1 week before the experimental trials, the experimenters introduced the robot in class during their daily “circle time”, as this provided a safe and familiar environment with the whole group in which the pedagogical assistants usually introduce new topics or new activities. One experimenter first introduced the robot by telling a story about Robin, the name of our robot, using a makeshift picture book. In this story we explained the similarities and dissimilarities between the robot and children to construct the type of common ground considered to have a positive effect on the learning outcome (Kanda et al., 2004). For example, we told that Robin enjoys dancing and wants to meet new friends, and even though he does not have a mouth and because of that cannot smile, he can smile using his eye LEDs.

After this story, another experimenter entered the room with the robot while it was actively looking at faces to provide an animate feeling. The robot introduced itself with a small story about itself and by performing a dance in which the children were encouraged to participate. The end of the circle time consisted of getting a blanket for the robot so it could “sleep”. This introduction was repeated later on the days we conducted the experiment in one-on-one sessions. While by then most children were comfortable interacting with the robot, some were still timid and anxious. To encourage these children to feel comfortable, one of the experiment leaders would sit next to the child during the warm-up phase of the experiment and motivate the child to respond to the robot when necessary until the child was sufficiently comfortable to interact with the robot by herself/himself. We found that the younger 3-year olds required more support from the experimenters than the older 3-year olds (Baxter et al., 2017). Although we are still analyzing the experiments, preliminary findings suggest that our introduction helped children to build trust and common ground with the robot effectively.

Temporal Contingency

Research has shown that it is crucial for children’s language development that their communication bids are responded to in a temporally contingent manner (Bornstein et al., 2008; McGillion et al., 2013). This, however, faces a technological challenge. While adults tend to take over turns very rapidly, robots require relatively long processing time to produce a response. Nevertheless, in our first experiment (de Haas et al., 2016), we observed that children were at first surprised by the delayed responses, but quickly adapted to the robot and waited patiently for a response. Perhaps this is because children also require longer than adults to take turns

(Garvey and Berninger, 1981) and having framed the robot as a peer children made the delays more plausible or expected. Nevertheless, while a lag in temporal contingency may not harm the interaction with children, it may harm learning. One way to remedy this may be to have the robot start responding by providing a backchannel signal, such as “uhm” to indicate the robot is (still) taking his turn, but requires more time to process (Clark, 1996).

Semantic Contingency

Robots should not only respond to children in a timely fashion, but also in a semantically contingent fashion (i.e., consistent with the child’s focus of attention), as this too has a positive effect on children’s language acquisition (Bornstein et al., 2008; McGillion et al., 2013). For instance, research has shown that by responding in a semantically contingent manner, either verbally or by following children’s gaze, (joint) attention is sustained for a longer duration (Yu and Smith, 2016), allowing children to learn more about a situation. To achieve semantically contingent responses, the robot should be able to understand the child’s communication bids, construct joint attention with the child, or at least identify what the child is attending to. Monitoring children’s behavior and establishing joint attention are therefore considered crucial for designing a successful robot tutor.

Monitoring Children’s Behavior

To understand children’s communication bids, as well as to test their pronunciation of the L2, it is important that the robot be equipped with well-functioning automatic speech recognition (ASR). However, the performance of state-of-the-art ASR for children is still suboptimal, especially for preschool-aged children (Fringi et al., 2015; Kennedy et al., 2017). Reasons for this include that children’s pronunciation is often flawed and that their speech has a different pitch than adults. Moreover, relatively little research has been carried out in this domain and not much data exist to train ASR on. While it can be expected that the performance of ASR for children will improve in the not too distant future (Liao et al., 2015), until then alternative strategies need to be developed that do not (exclusively) rely on ASR.

In our project, we explore various strategies to achieve this, both based on monitoring non-verbal behaviors of the children and focusing on comprehending rather than producing L2. The first strategy relies on providing children tasks they have to perform in the learning environment, such as placing “a toy cow behind a tree” when teaching spatial language. This, however, requires the visual object recognition on the robot to work well, which is only the case when the scene contains a limited set of distinctively recognizable objects, such as distinctly colored objects (Nguyen et al., 2015). A potential solution explored in our project is to use objects with build-in RFID sensors that can be tracked automatically. The second solution we explore is to use a touch screen tablet that displays scenes the child can manipulate, which not only has the advantage of avoiding the problem of object recognition, but also allows us to control

the robot's responses and vary the scenes in real time. A downside, however, is that it takes away the 3-dimensional physical aspect of embodied cognition that would help the children to better entrench what they learn (Glenberg, 2010). Currently, experiments are underway to investigate the effect of using real vs. virtual objects. These solutions not only aid in understanding the child's communication bids, it also helps in identifying their attention and can thus contribute to establishing joint attention.

Joint Attention and Gestures

Joint attention, where interlocutors attend on the same referent, is a form of social interaction that has been shown to support children's language learning (Tomasello and Farrar, 1986). One way to establish joint attention with a child is to guide their attention to a referent using gestures, such as pointing or iconic gestures. The ability to produce gestures in the real world is potentially one of the main advantages of using physical robots as opposed to virtual agents, who may have a harder time to establish joint attention. However, many robots' physical morphologies do not correspond one-to-one to the human body. Hence, many human gestures cannot be translated directly to robot gestures. For instance, the NAO robot that we use in our research has a hand with three fingers that cannot be controlled independently, so index finger pointing cannot be

achieved (see **Figure 1**). Will children still recognize NAO's arm extension as a pointing gesture? And if so, will they be able to identify the object the robot refers to? We are currently running an experiment to investigate how NAO's pointing gestures are perceived, and preliminary findings show that participants have difficulty identifying the referred object on a small tablet screen. Similar issues arise when developing other gestures. One of the other non-verbal behaviors we are using is the coloring of NAO's eye LEDs to indicate the robot's happiness as a form of positive feedback, since the robot cannot smile with its mouth.

Feedback

Feedback, too, is an interactional feature known to help language learning (Matthews et al., 2007; Ateş-Şen and Küntay, 2015). The question is how should the robot provide feedback, such that it is both pleasant and effective for learning? While adults provide positive feedback explicitly, they usually provide negative feedback implicitly by reformulating children's errors in the correct form. In child-child interactions, however, Long (2006) found that there was a clear advantage in learning from explicit negative feedback (e.g., by saying "no, that's wrong, you need to say 'he ran'") when compared to reformulating feedback (the learner says "he runned" and the teacher reacts with "he ran").



FIGURE 1 | NAO pointing to a block with three fingers. (Note that written, informed consent was obtained from the parents of the child for the publication of this image).

To investigate how children experience feedback from a peer robot, we carried out an experiment among 85 3-year-old Dutch-speaking children at preschools in Netherlands (de Haas et al., 2016, 2017). In this experiment, the children interacted with a NAO robot during which they received a short lesson on how to count from 1 to 4 in English. After a short training phase, in which the children were presented with the four counting words twice in relation to body parts and wooden blocks, they were given instructions by the robot to pick up a given number of blocks. While the instructions were given in their native language, the numbers were uttered in English. In response to the child's ability to achieve the task, the robot provided feedback. The experiment followed a between-subjects design with three conditions: adult-like feedback (explicit positive and implicit negative), peer-like feedback (no positive and explicit negative) and no feedback. We did not find significant differences in learning gain between the conditions, probably because the target words were insufficiently often repeated. However, we explored the way in which the children engaged with the robot after they received feedback and we found that children looked less often at the experimenter in the feedback conditions than in the no feedback condition. Further analyses are carried out to evaluate how the children responded to the various forms of feedback to find out what type of feedback would be most effective for achieving both acceptable and effective tutoring interactions.

Zone of Proximity and Adaptivity

Finally, from a pedagogical point of view it is desirable that the interactions between child and robot be sufficiently challenging and varied so that the child has a target to learn from, but at the same time interactions should not be too difficult, because that may frustrate the child causing it to lose interest in the robot (Charisi et al., 2016). In other words, the robot should remain in Vygotsky's Zone of Proximity that supports an effective learning environment (Vygotsky, 1978). In order to achieve this, the robot should be able to keep track of the children's advancements in language learning and perhaps their emotional states during the tutoring sessions, and adapt to these. While the former can be monitored as discussed previously, it may be possible to detect emotional states known to influence learning (e.g., concentration, confusion, frustration and boredom) using methods from affective computing (D'Mello and Graesser, 2012). Using this type of information, it is possible to adapt the tutoring sessions by either reducing or increasing the number of repetitions, and/or change the subject (Schodde et al., 2017).

CONCLUSION

This perspective article presented some design features that we consider crucial for developing a social robot as an effective second language tutor. We believe the robot is most effective when it is framed as a peer, i.e., as a fellow language learner and playmate, but that is designed to use adult-like interaction strategies to optimize learning efficacy. In order to establish common ground and trust to facilitate long-term interactions, we consider it essential that the robot be introduced with appropriate care on the first encounter. As an example, we

outlined our strategy for introducing a robot to preschool children. Interactions between child and robot should be contingent and multimodal, and provide appropriate forms of feedback. We argued that the robot should remain within Vygotsky (1978) Zone of Proximal Development and thus should adapt to the individual level of the child.

We also discussed some technical challenges that need to be solved in order to implement contingent interactions; the most important of which we believe is ASR, which presently does not work well for children's speech. While various technical challenges still remain, we expect that social robots will provide effective digital technologies to support second language development in the years to come.

The present list of design features covers many aspects that need to be considered when developing a tutor robot, but it is not yet comprehensive. One aspect that has not been covered, for instance, concerns the design of robots for children from different cultures, which could require different design choices (Shahid et al., 2014). For example, in some cultures education is more teaching-centered (Hofstede, 1986) and thus designing the tutor as a peer robot may be less effective or acceptable (Tazhigaliyeva et al., 2016). Concluding, this perspective article offers only a first step towards a comprehensive list of design features for tutor robots and additional research is needed to complete and optimize the list.

ETHICS STATEMENT

The Research Ethics Committee of Tilburg School of Humanities approved this study, and the parents of all participating children gave written informed consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

PV, MH and EK designed the conceptual aspects of the article; PV, MH, CJ and PB carried out the literature review; PV, EK and MH designed the feedback study; MH, CJ and PB designed the introduction study; MH, CJ and PB carried out the studies; PV and MH wrote the article; CJ, PB and EK revised the article critically.

FUNDING

This work has been supported by the EU H2020 L2TOR project (grant 688014). CJ and PB thank the research trainee program of the Tilburg School of Humanities for their support.

ACKNOWLEDGMENTS

The authors wish to thank all members of the L2TOR project for their support and advice regarding this research. We also thank Kinderopvanggroep Tilburg and all participating daycare centers and preschools for their assistance in this research. Finally, a big thank you to all the children and their parents for participating in our research.

REFERENCES

- Ateş-Şen, B. A., and Küntay, A. C. (2015). Children's sensitivity to caregiver cues and the role of adult feedback in the development of referential communication. *The Acquisition of Reference*, eds L. Serratrice and S. E. M. Allen (Amsterdam: John Benjamins), 241–262.
- Baxter, P., De Jong, C., Aarts, A., de Haas, M., and Vogt, P. (2017). "The effect of age on engagement in preschoolers' child-robot interactions," in *Companion proceedings of the 12th Annual ACM International Conference on Human-Robot Interaction (HRI'17)*, (Vienna, Austria).
- Belpaeme, T., Kennedy, J., Baxter, P., Vogt, P., Krahmer, E. E. J., Kopp, S., et al. (2015). "L2TOR-second language tutoring using social robots," in *First Workshop on Educational Robots (WONDER)*, (Paris, France).
- Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C. S., and Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: longitudinal analysis of a multidimensional, modular and specific parenting construct. *Dev. Psychol.* 44, 867–874. doi: 10.1037/0012-1649.44.3.867
- Charisi, V., Davison, D., Reidsma, D., and Evers, V. (2016). "Children and robots: a preliminary review of methodological approaches in learning settings," in *2nd Workshop on Evaluating Child Robot Interaction - HRI*, (Christchurch, New Zealand).
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human-robot interaction. *Philos. Trans. R. Soc. B Biol. Sci.* 1480, 679–704. doi: 10.1098/rstb.2006.2004
- de Haas, M., Baxter, P., de Jong, C., Vogt, P., and Krahmer, E. (2017). "Exploring different types of feedback in preschooler and robot interaction," in *Companion proceedings of the 12th Annual ACM International Conference on Human-Robot Interaction (HRI'17)*, (Vienna, Austria).
- de Haas, M., Vogt, P., and Krahmer, E. J. (2016). "Enhancing child-robot tutoring interactions with appropriate feedback," in *Proceedings of First Workshop on Long-Term Child-Robot Interaction. IEEE Ro-Man*, (New York, NY).
- D'Mello, S., and Graesser, A. (2012). Dynamics of affective states during complex learning. *Learn. Instr.* 22, 145–157. doi: 10.1016/j.learninstruc.2011.10.001
- Fridin, M. (2014). Kindergarten social assistive robot: first meeting and ethical issues. *Comput. Hum. Behav.* 30, 262–272. doi: 10.1016/j.chb.2013.09.005
- Fringi, E., Lehman, J., and Russell, M. J. (2015). "Evidence of phonological processes in automatic recognition of children's speech," in *16th Annual Conference of the International Speech Communication Association*, (Dresden, Germany), 1621–1624.
- Garvey, C., and Berninger, G. (1981). Timing and turn taking in children's conversations. *Discourse Process.* 4, 27–57. doi: 10.1080/01638538109544505
- Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdiscip. Rev. Cogn. Sci.* 4, 586–596. doi: 10.1002/wcs.55
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., and Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Comput. Assist. Lang. Learn.* 27, 70–105. doi: 10.1080/09588221.2012.700315
- Han, J. H., Jo, M. H., Jones, V., and Jo, J. H. (2008). Comparative study on the educational use of home robots for children. *J. Inf. Process. Syst.* 4, 159–168. doi: 10.3745/jips.2008.4.4.159
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., and Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Hum. Factors* 53, 517–527. doi: 10.1177/0018720811417254
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Dev. Psychol.* 49, 4–14. doi: 10.1037/a0027238
- Hofstede, G. (1986). Cultural differences in teaching and learning. *Int. J. Intercult. Relat.* 10, 301–320. doi: 10.1016/0147-1767(86)90015-5
- Hood, D., Lemaignan, S., and Dillenbourg, P. (2015). "When children teach a robot to write: an autonomous teachable humanoid which uses simulated handwriting," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, (New York, NY), 83–90.
- Hsin, C.-T., Li, M.-C., and Tsai, C.-C. (2014). The influence of young children's use of technology on their learning: *A. Educ. Technol. Soc.* 17, 85–99. Available online at: <https://eric.ed.gov/?id=EJ1045554>
- Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: a field trial. *Hum. Comput. Interact.* 19, 61–84. doi: 10.1207/s15327051hci1901&2_4
- Kennedy, J., Baxter, P., Senft, E., and Belpaeme, T. (2015). "Higher nonverbal immediacy leads to greater learning gains in child-robot tutoring interactions," in *International Conference on Social Robotics*, eds A. Tapus, E. André, J.-C. Martin, F. Ferland and M. Ammi (New York, NY: Springer International Publishing), 327–336.
- Kennedy, J., Baxter, P., Senft, E., and Belpaeme, T. (2016). "Social robot tutoring for child second language learning," in *Proceedings of the 11th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI'16)*, (Christchurch, New Zealand), 231–238.
- Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F., et al. (2017). "Child speech recognition in human-robot interaction: evaluations and recommendations," in *Proceedings of the 12th Annual ACM International Conference on Human-Robot Interaction (HRI'17)*, (Vienna, Austria).
- Leyzberg, D., Spaulding, S., Toneva, M., and Scassellati, B. (2012). "The physical presence of a robot tutor increases cognitive learning gains," in *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, (Sapporo, Japan), 1882–1887.
- Liao, H., Pundak, G., Siohan, O., Carroll, M. K., Coccaro, N., Jiang, Q. M., et al. (2015). "Large vocabulary automatic speech recognition for children," in *Interspeech*, (Dresden, Germany), 1611–1615.
- Long, M. H. Ed. (2006). "Recasts in SLA: the story so far," in *Problems in SLA. Second Language Acquisition Research Series*, (Mahwah, NJ: Lawrence Erlbaum Associates), 75–116.
- Lund, H. H. (2003). "Adaptive robotics in the entertainment industry," in *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation (cira2003)*, (Vol. 2), (Kobe, Japan), 595–602.
- Mashburn, A. J., Justice, L. M., Downer, J. T., and Pianta, R. C. (2009). Peer effects on children's language achievement during pre-kindergarten. *Child Dev.* 80, 686–702. doi: 10.1111/j.1467-8624.2009.01291.x
- Matthews, D., Lieven, E., and Tomasello, M. (2007). How toddlers and preschoolers learn to uniquely identify referents for others: a training study. *Child Dev.* 6, 1744–1759. doi: 10.1111/j.1467-8624.2007.01098.x
- McGillion, M., Herbert, J., Pine, J., Keren-Portnoy, T., Vihman, M., and Matthews, D. (2013). Supporting early vocabulary development: what sort of responsiveness matters? *IEEE Trans. Auton. Ment. Dev.* 5, 240–248. doi: 10.1109/tamd.2013.2275949
- Nguyen, T. L., Boukezzoula, R., Coquin, D., Benoit, E., and Perrin, S. (2015). "Interaction between humans, NAO robot and multiple cameras for colored objects recognition using information fusion," in *8th International Conference on Human System Interaction (HSI)*, (Warsaw, Poland), 322–328.
- Schodde, T., Bergmann, K., and Kopp, S. (2017). "Adaptive robot language tutoring based on bayesian knowledge tracing and predictive decision-making," in *Proceedings of the 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2017)*, (Vienna, Austria).
- Shahid, S., Krahmer, E., and Swerts, M. (2014). "Child-robot interaction across cultures: how does playing a game with a social robot compare to playing a game alone or with a friend? *Comput. Human Behav.* 40, 86–100. doi: 10.1016/j.chb.2014.07.043
- Tanaka, F., and Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: field experiments in a classroom for vocabulary learning. *J. Hum. Robot Interact.* 1, 78–95. doi: 10.5898/jhri.1.1.tanaka

- Tazhigaliyeva, N., Diyas, Y., Brakk, D., Aimambetov, Y., and Sandygulova, A. (2016). "Learning with or from the robot: exploring robot roles in educational context with children," in *International Conference on Social Robotics*, eds A. Agah, J.-J. Cabibihan, A. M. Howard, M. A. Salichs and H. He (New York, NY: Springer International Publishing), 327–336.
- Tomasello, M., and Farrar, M. J. (1986). Joint attention and early language. *Child Dev.* 57, 1454–1463. doi: 10.2307/1130423
- Vygotsky, L. (1978). *Mind in Society*. Harvard: Harvard University Press.
- Westlund, J. M. K., Martinez, M., Archie, M., Das, M., and Breazeal, C. (2016). "Effects of framing a robot as a social agent or as a machine on children's social behavior," in *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, eds S. Y. Okita, T. Shibata and B. Mutlu (Washington, DC: IEEE), 688–693.
- Yu, C., and Smith, L. B. (2016). The social origins of sustained attention in one-year-old human infants. *Curr. Biol.* 26, 1235–1240. doi: 10.1016/j.cub.2016.03.026

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Vogt, de Haas, de Jong, Baxter and Kraemer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.