Reducing communication failures in robot education

A Preprint

Ronald Cumbal
KTH - Royal Institute of Technology
Stockholm - Sweden
ronaldcg@kth.se

Olov Engwall
KTH - Royal Institute of Technology
Stockholm - Sweden
engwall@kth.se

José Lopes
Heriot-Watt University
Edinburgh - United Kingdom
jd.lopes@hw.ac.uk

February 18, 2021

Abstract

Social robots have shown promising to support both the learning process and the subjective learning experience in educational tasks. However, limitations on the system's components (e.g., speech recognition errors) limits their potential. In this paper, we propose a method that leverages the learner to correct robot mistakes in a second language learning activity with a social robot. This approach takes advantage from the participation of the learners to correct any probable robot mistakes. We describe the proposed study as part of a second language learning practice with a social robot. We evaluate what are the implications from the user perspective and the efficacy of the proposed method in educational interactions.

***K*eywords** Second language learning · robot failure · conversation practice

1 Introduction

Educational robots can generate positive effects in both the learning outcome and the affective state of a learner [1]. However, failure of system components, such as automatic speech recognition or emotion classification, have limited the potential of such robots [1]. A common solution to reduce these errors employs human controllers to manage different aspects of the robot behavior, i.e. Wizard-of-Oz setups, but this option is less reproducible and removes the expected autonomy of a robot. Alternatively, researchers often constrain the level of interaction or means of communication, e.g., using screens or buttons to collect user responses, to minimize communication errors. However, such decision may limit the naturalness of the interaction.

The solution for these inconveniences is non-trivial, as social robots in education need to make decisions that support the learning process, in addition to sustaining an engaging interaction. Inappropriate behaviors, e.g. timing actions, may generate an adverse effects, like distracting the student [2, 3]). Therefore, in educational settings, robot behaviors require robust implementations that may surpass the common state-of-the-art results in interactive technology. Current solutions, like fine-tuning automatic speech recognizer systems to the intended users, are often practical remedies, but not efficient or error-proof.

We propose a different approach that could reduce communication errors with educational robots. We use a multi-learner interaction (i.e. more than one learner in addition to the robot) that enables the robot to take advantage of the learners by redirecting an action when it is confused on what decision to take. For example, the robot may refrain from answering a question when uncertain about the response, and instead prompt another learner to intervene. Furthermore, if the robot is uncertain about the state of the interaction, it may also request a non-active learner to evaluate the robot's next actions, e.g., *"would you say A or B now?"*, where A and B corresponds to the best ranked next dialogue acts. This behavior can be both direct and indirect. For example, instead of verbally asking a learner, it could use gestures or other non-verbal expressions to hint a certain request.

2 Background and Research Questions

Robot failure in human-robot interaction has been extensively studied in previous works, for a comprehensive review we suggest to read [4]. However, the implication of robot failures in educational environments has attracted less interest. We can assume that when the robot takes the role of a tutor, there is a higher expectation on it's functionality, as shown by Alves-Oliviera *et. al.* [5]. In that study, children expressed a high expectation on the robot's capabilities, but also indicated that these expectations were satisfied. However, the system was partially controlled by a human and all communication was realized through a touch screen. Hence, communication errors may not have influenced the outcome. On the other hand, when the robot assumes a different role, this expectation may be less rigorous. For example, Takana *et. al.* [6] showed that when a robot is presented as a peer, the robot could be expected to make mistakes and these errors could actually be used to encourage learning. Nonetheless, more research is required to better understand the effect of unexpected mistakes in human-robot interaction with educational content.

In particular, the aspect of verbal communication is an important element for social robots designed for education and requires careful design to avoid unintentional expectations [4]. Cha *et. al.* found that robots who displayed conversational skills were perceived as less capable than robots that only use functional speaking after they made a mistake [7]. However, as mentioned before, simple verbal commands may sometimes limit the natural interaction that can develop in educational settings. For example, learners are often interested in additional information or query for explanations. Therefore, while a full conversational system is not suggested, educational robots should attempt to develop interactions that extend to more than just functional commands.

With these considerations, an alternative approach to avoid communication mistakes may come from allowing the robot to elicit a reaction from one learner when the robot is uncertain. In educational settings, this option could be build under the paradigm of constructive education. In this process, active participation is necessary and forms the means by which a learner acquires knowledge [8]. Furthermore, considering the current tendencies to create collaborative environments [9], this alternative would just boost cooperation among learners. Our assumption, then, implies that the solicited action for the learner should be considered another form of participation, even if the learner is not ready to intervene. Related work on *cold-calling* students during a discussion (i.e., unpredictably getting called to answer a question) has shown that these actions do not make students feel uncomfortable and actually generate more participation over time [10]. A particular benefit is also found from the easiness in which learners share idioms to communicate effectively, which does not translate to interactions with tutors [11].

Building on this background, we will attempt to explore the following research questions:

- Q1: Can verbal communication errors be reduced in robot assisted education by eliciting the intervention of a non-active student when the robot is uncertain on what to say?
- Q2: What is the perception of a student that interacts with a robot that solicits intervention from another peer when it is uncertain (i.e. effects on affective behavior)?

3 Design

In order to explain the method we present in this paper, we employ the next two sample cases. The first scenario, shown in Table 1, evolves from a practice conversation developed between a social robot and two second language learners [12]. The second case, shown in Table 2, resembles interactions in a vocabulary practice game between a social robot, a native speaker and a second language learner.

Table 1: Sample case: practice conversation in personal topics.

Subject	Verbal Interaction	Robot State	Robot Action
...			
Learner 1	I enjoy listening to <*music*>	Positive utterance	Use positive expression
Robot	That's very interesting.		
Learner 1	Like <*homeland artist*>, do you *now*?	Uncertain, *ASR confidence*	Other learner intervene
Robot	What do you think *Learner 2*?		
Learner 2	I don't know them, but I like <*music*> too!	Positive utterance	Use positive expression
Robot	Very nice, tell me more.		
...			

Table 2: Sample case: practice vocabulary word (i.e. *baseball*) in a guessing game.

Subject	Interaction	Robot state	Intervention Required
Learner 1	This is a sport played with a ball.	Guess <*sport*>	Say guess
Robot	Is it *soccer*?		
Learner 1	No, I played it yesterday with <*not recognized*>	Uncertain, *ASR confidence*	Other learner intervene
Robot	What do you think *Learner 2*?		
Learner 2	Yes, a ball and a bat	Guess <*sport*>	Say guess
Robot	Is it *baseball*?		

3.1 What is the context to your research?

We center our research on language learning. As this setting is heavily dependent on verbal communications, we can evaluate the proposed approach as a holistic solution. Research in robot assisted language learning (RALL) has increased considerably over the last decade [13]. Nonetheless, the performance of automatic speech recognition systems can easily fail to produce accurate results when the user population differs from the one used to develop the technology, e.g., children [14]. Consequently, it is highly probable that interactions with non-native learners will cause communication errors. In our proposed study we will employ a vocabulary practice game and a conversation practice setup for Swedish language learners. In both scenarios, the robot takes the role of a native speaker, but not necessarily a tutor.

3.2 Who are your learners?

Our work focuses on young-adults or adult learners of Swedish second language. In order to produce a minimal progression of spoken interactions or short conversations, the language learners need to be at intermediate level of speaking proficiency. Furthermore, it is preferred that learners don't surpass an advanced level, as the practice activity may be less engaging for them [12]. Therefore, according to the Common European Framework of Reference for Languages (CEFR) [15], the participants should have a minimum Swedish proficiency level of A2 (low-intermediate), but no higher than C2 (advanced).

3.3 What are the learning objectives?

The outcome of vocabulary or speaking practice can be measured as an improvement in (spoken) language production, e.i., the amount of coherent speech that a learner generates on its own over a period of time. Intuitively, the new vocabulary words should be part of this speech. From this perspective, the proposed method (i.e., to avoid communication mistakes) should enable the learners to practice more words and expressions in conversation that does is not easily interrupted by systematic mistakes.

3.4 Where is the learning occurring? (home, school, elderly facility ...)

The setup of these experiments is not limited to classical classroom environments. However, as these are designed to include at least two learners, we can assume that the experiments will take place in educational centers.

3.5 What robot do you use and why?

We will use the anthropomorphic robot-head *Furhat* [16]. This robot platform can be used to develop spoken interactions with multiple users while incorporating other forms of non-verbal communication, e.g., nodding, back-channels or facial expressions.

4 Assessment

4.1 What methods do you plan to use for the assessment?

A clear evaluation of language production is hard to define for practice activities. A practical solution, in the case of the conversational practice, can be formulated in similar terms to the evaluation metrics of dialogues systems. Therefore, the practice conversation could be measured on the mean length of utterances (MLU), the total length of utterances, and the number of turns in the complete conversation. Furthermore, vocabulary practice can be measured through the correct use of the intended vocabulary. We can use standard pre- and post-test to evaluate vocabulary retention or measure their use during the conversation. Finally, to evaluate user perspectives we can employ a modified version of the "Goodspeed" questionnaire or similar methods.

4.2 How do they map with the learning objectives?

It was hard to define this mapping, as evaluating the learning performance is not an easy task for practice conversations or vocabulary practice. We hope some discussion among peers attending the workshop would be able to shine some light on this manner.

Acknowledgments

We would like to thank Sanne van Waveren for her revision of this article and support on developing the idea.

References

[1] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. Social robots for education: A review. *Science Robotics*, 3(21), 2018.

[2] James Kennedy, Paul Baxter, and Tony Belpaeme. The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 67–74. IEEE, 2015.

[3] Elmira Yadollahi, Wafa Johal, Ana Paiva, and Pierre Dillenbourg. When deictic gestures in a robot can harm child-robot collaboration. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*, IDC '18, page 195–206, New York, NY, USA, 2018. Association for Computing Machinery.

[4] Shanee Honig and Tal Oron-Gilad. Understanding and resolving failures in human-robot interaction: Literature review and model development. *Frontiers in Psychology*, 9:861, 2018.

[5] Patrícia Alves-Oliveira, Tiago Ribeiro, Sofia Petisca, Eugenio di Tullio, Francisco S. Melo, and Ana Paiva. An empathic robotic tutor for school classrooms: Considering expectation and satisfaction of children as end-users. In Adriana Tapus, Elisabeth André, Jean-Claude Martin, François Ferland, and Mehdi Ammi, editors, *Social Robotics*, pages 21–30, Cham, 2015. Springer International Publishing.

[6] Fumihide Tanaka and Shizuko Matsuzoe. Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *J. Hum.-Robot Interact.*, 1(1):78–95, July 2012.

[7] Elizabeth Cha, A. Dragan, and S. Srinivasa. Perceived robot capability. *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 541–548, 2015.

[8] Sigmund Tobias and Thomas M Duffy. *Constructivist instruction: Success or failure?* Routledge, 2009.

[9] Ido Roll and Ruth Wylie. Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26, 06 2016.

[10] Elise J. Dallimore, Julie H. Hertenstein, and Marjorie B. Platt. Impact of cold-calling on student voluntary participation. *Journal of Management Education*, 37(3):305–341, 2013.

[11] Roland C. Christensen and David A. Garvin. Every student teaches and every teacher learns: The reciprocal gift of discussion teaching. In Ann Sweet, editor, *Education for judgment: The artistry of discussion leadership*. Harvard Business School Press, Boston, MA, 1991.

[12] Olov Engwall, José Lopes, and Anna Åhlund. Robot interaction styles for conversation practice in second language learning. *International Journal of Social Robotics*, 2020.

[13] Natasha Randall. A survey of robot-assisted language learning (rall). *J. Hum.-Robot Interact.*, 9(1), December 2019.

[14] James Kennedy, Séverin Lemaignan, Caroline Montassier, Pauline Lavalade, Bahar Irfan, Fotios Papadopoulos, Emmanuel Senft, and Tony Belpaeme. Child speech recognition in human-robot interaction: Evaluations and recommendations. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '17, page 82–90, New York, NY, USA, 2017. Association for Computing Machinery.

[15] Council of Europe. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment : Case Studies*. Language Learning. Council of Europe Publishing/Éditions du Conseil de l'Europe, 2002.

[16] Samer Al Moubayed, Jonas Beskow, Gabriel Skantze, and Björn Granström. Furhat : A back-projected human-like robot head for multiparty human-machine interaction. In *Cognitive Behavioural Systems : COST 2102 International Training School, Dresden, Germany, February 21-26, 2011, Revised Selected Papers*, number 7403 in Lecture Notes in Computer Science, pages 114–130, 2012. QC 20121123.