Second Language Tutoring using Social Robots

Project No. 688014

L2TOR

Second Language Tutoring using Social Robots

Grant Agreement Type: Collaborative Project
Grant Agreement Number: 688014

D7.5 Evaluation report of L2TOR versus tablet

Due Date: **31/12/2018**
Submission Date: **21/01/2019**

Start date of project: **01/01/2016** Duration: **36 months**

Organisation name of lead contractor for this deliverable: **TIU**

Responsible Person: **Paul Vogt** Revision: **1.0**

Project co-funded by the European Commission within the H2020 Framework Programme		
Dissemination Level		
PU	Public	**PU**
PP	Restricted to other programme participants (including the Commission Service)	
RE	Restricted to a group specified by the consortium (including the Commission Service)	
CO	Confidential, only for members of the consortium (including the Commission Service)	

Contents

Executive Summary

This deliverable discusses the results of the tablet-only and robot conditions of the large-scale field study. The design and the results of the evaluation study have been reported in previous deliverables. We therefore only will briefly summarise relevant methods and results for this deliverable, and discuss the findings in some more detail. The findings show that children's learning gains were not different between the robot and tablet-only conditions. In this deliverable we argue that the lack of differences between tablet and robot conditions may be due to the prominent role of the tablet game in all conditions. As a result, the robot might distract the children from their learning tasks during the lessons. Finally, we provide recommendations for the future.

Principal Contributors

TIU: Mirjam de Haas, Emiel Krahmer, Paul Vogt, Bram Willemsen, Jan de Wit
UU : Rianne van den Berghe

Revision History

Version 1.0 (PV 21-01-2019)
This is the first version.

1 Introduction

This deliverable reports on the results of task T7.2 (**Compare L2TOR with tablet-based digital learning environment**). The objective for this comparison is to investigate to what extent and in what ways the robot in the L2TOR system has an added value with respect to a tablet-based learning environment. This is important for evaluating the benefits of social robots for second language tutoring.

To carry out this comparison, we included a 'tablet only' condition to our large-scale field study as explained in deliverables D7.2, D7.6 and in Vogt et al. (2019). As described extensively in deliverables D7.1 and D7.2, the children that participated in the tablet-only condition received the same lessons as the children participating in the two robot conditions. Except that in the tablet-only condition, the robot was not visibly present and its voice came from the tablet's loudspeakers. While the robot was hidden, we still used the robot's software in the experiment. Moreover, we adapted the script of what the robot (and tablet) would say for all conditions to account for the fact that the robot was not present. This was done so that the story would be consistent with the setting (see D2.3).

As already presented in D7.2 and D7.6, the evaluations carried out so far, did not reveal any significant differences in learning outcomes between the two robot conditions and the tablet-only condition (see D7.2; Vogt et al., 2019). Table 1 summarizes these results.

Table 1. Summary of the main test results from the large field study, comparing the four conditions concerning the translation tasks (English-Dutch and Dutch-English), and the comprehension scores during the pretest, post-test and delayed (retention) post-test. Reprinted from Vogt et al. (2019).

Condition / Test	Pre-test	Post-test	Retention
Iconic gesture			
Trans(En-Du)	3.38 (3.07)	7.47 (5.16)	8.15 (5.01)
Trans(Du-En)		6.08 (4.19)	6.57 (4.65)
Comprehension		29.30 (5.80)	30.45 (6.29)
No iconic gesture			
Trans(En-Du)	3.59 (3.14)	7.83 (4.94)	8.02 (4.92)
Trans(Du-En)		6.54 (4.28)	6.44 (4.59)
Comprehension		29.50 (6.13)	30.45 (6.29)
Tablet only			
Trans(En-Du)	3.91 (2.80)	7.70 (4.73)	8.42 (4.75)
Trans(Du-En)		6.49 (4.10)	6.70 (4.29)
Comprehension		29.38 (6.44)	30.17 (6.60)
Control			
Trans(En-Du)	2.81 (2.83)	3.81 (3.21)	4.34 (3.22)
Trans(Du-En)		3.16 (2.27)	3.47 (2.13)
Comprehension		25.03 (6.66)	26 (6.04)

All scores indicate the average number of words correctly translated or comprehended. Minimum scores are 0, maximum scores are 34 for translation and 54 for comprehension. For comprehension, chance level is 18.

2. Interpretation

To understand why we find these similarities in the results between the three conditions, it is instructive to look at the similarities in the experimental setup between the conditions.

Role of the tablet

While we could not avoid the use of the tablet in the lessons, it is likely that the children's learning greatly depended on the tablet game. The tablet provided different types of input during the lesson series in all conditions:

1. The contextual setting of all lessons was provided through the tablet. Children had to pay attention to the tablet to understand the lessons.
2. Children were instructed to perform various tasks on the tablet (touching and moving objects).
3. New target words in L2 were introduced by a human voice through the tablet.
4. The interactive story with instructions, translations and feedback as spoken by the robot in the two robot conditions was the same as in the tablet condition and delivered by the same voice. The only differences were that in the robot conditions, the robot was physically present, and delivered the story using gestures (the amount and type of which varied between the two robot conditions) and some additional social cues (e.g., mutual eye-gaze and non-verbal feedback by colouring the eyes).

Robot might distract the children from the lesson material

While various studies have shown positive effects on learning regarding the use of (iconic) gestures (de Wit et al., 2018) and the mere presence of the robot compared to tablet (Kennedy et al., 2017), our long-term field study did not reveal this. It thus seems that the potential benefits of robots in these tutoring sessions may have been cancelled by the disadvantages incorporated in our experimental setup. We actually believe that the robot may have distracted children from playing the lesson game on the tablet. Most of what the children needed to do was carrying out tasks on the tablet and paying attention to things happening on the tablet. In the tablet condition, children could remain focused on the tablet, while in the robot conditions, the children had to alternate their focus of attention between the tablet and robot. So, if there were to be an advantage of tutoring by the social robot in our setup, then this advantage could have been cancelled by the advantage of the tablet only condition for not having to switch attention between tablet and robot.

In a sense, this interpretation seems to be confirmed by the preliminary results on the analysis of engagement. Although we only have data from the first two lessons on engagement, we see that the observers rated the children higher on task engagement in the tablet-only condition than in the robot conditions. Children in the tablet-only condition did not have to focus on the robot, and could thus focus more on the tablet game (hence were more engaged with the tablet game). As already mentioned in D7.6, if this pattern remains for the other lessons and for all children, this will support our interpretation. We have not yet carried out any statistical analyses on the engagement data, as not enough videos have been rated yet for the tablet-only condition (this is still work in progress).

3. Conclusion

It is important to stress that we did not find a significant difference in learning gain between the robot conditions and tablet-only condition in this particular setup. This does not necessarily mean that robots cannot have an added value in second language tutoring compared to a tablet application. In the present model, the tablet had a central role for delivering content and controlling the interaction. The reason for having the tablet occupying such a central role was induced by our objective of designing an autonomous tutoring system. Technological limitations (esp. ASR for child speech and automatic object recognition) prevented us from managing verbal-driven interactions between child and robot. While in future applications the tablet could still play a role, we believe it is best to reduce its role as much as possible to create added value of the robot. The tablet could, for example, be used to provide children or teachers with choices as to what to learn, while the actual games are played with a robot in which its tactile sensors could fulfil a role in managing turn-taking. Also, additional external sensors could be used for measuring social cues. This could be touch-based sensors or microphones worn by the child. However, for using sound reliably, the technology of ASR or automatic voice detection should be improved.

When more ubiquitous technologies can be used to sense children's social cues and keep track of both human behaviour and objects, more interactive games can be designed in which the robot's embodiment and situatedness of the interactions play a more central role. We believe that this would benefit language learning, since it allows for language learning in a physically grounded setting, something that –in principle– a tablet-only application cannot provide (at least not to the same extent). Future research should therefore concentrate, among others, on finding ways to incorporate ubiquitous sensing technologies for having more naturalistic child-robot interactions.

References

Kennedy, J., Baxter, P. & Belpaeme, T. (2017). The Impact of Robot Tutor Nonverbal Social Behaviour on Child Learning. Frontiers in ICT, 4, 6.

Vogt, P., van den Berghe, R., de Haas, M., Hoffmann, L., Kanero, J., Mamus, E., Montanier, J.-M., Oranç, C., Oudgenoeg-Paz, O., Papadopoulos, F., Schodde, T., Verhagen, J., Wallbridge, C.D., Willemsen, B., de Wit, J., Belpaeme, T., Bergmann, K., Göksun, T., Kopp, S., Krahmer, E., Küntay, A.C., Leseman & Pandey, A.K., "Second language tutoring using social robots: A large-scale study," to appear in the *2019 ACM/IEEE International Conference on Human-Robot Interaction*, 2019.

de Wit, J., Schodde, T., Willemsen, B., Bergmann, K., de Haas, M., Kopp, S., Krahmer, E. & Vogt, P (2018). The Effect of a Robot's Gestures and Adaptive Tutoring on Children's Acquisition of Second Language Vocabularies. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18). ACM, Chicago, IL, USA, 50-58