L2TOR - Second Language Tutoring using Social Robots

Tony Belpaeme[1], James Kennedy[1], Paul Baxter[1], Paul Vogt[2], Emiel E.J. Krahmer[2], Stefan Kopp[3], Kirsten Bergmann[3], Paul Leseman[4], Aylin C. Küntay[5], Tilbe Göksun[5], Amit K. Pandey[6], Rodolphe Gelin[6], Petra Koudelkova[6], Tommy Deblieck[7]

[1] Plymouth University, UK, [2] Tilburg University, The Netherlands, [3] University of Bielefeld, Germany, [4] University of Utrecht, The Netherlands, [5] Koç University, Turkey, [6] Aldebaran Robotics, France and [7] QBMT, Belgium

Abstract. This paper introduces a research effort to develop and evaluate social robots for second language tutoring in early childhood. The L2TOR project will capitalise on recent observations in which social robots have been shown to have marked benefits over screen-based technologies in education, both in terms of learning outcomes and motivation. As language acquisition benefits from early, personalised and interactive tutoring, current language tutoring delivery is often ill-equipped to deal with this: classroom resources are at present inadequate to offer one-to-one tutoring with (near) native speakers in educational and home contexts. L2TOR will address this by furthering the science and technology of language tutoring robots. This document describes the main research strands and expected outcomes of the project.

1 Background

Second language learning has become an important element of formal education for many children in Europe and beyond. For some children, the language used at school is a second language (noted as L2), as they speak a different language or dialect at home. This not only holds for immigrant children, but also for children speaking an official minority language of their country of residence. Preschool years are important to develop adequate knowledge of the academic language, as later educational success builds on it (Leseman & van den Boom, 1999; Hoff, 2013). Thus, it is essential that children with a different home language than the dominant one receive “sensitive” bilingual input and interaction once they enter day care and preschool settings. The robot tutor we propose here serves that crucial aim.

The current challenges of standard L2 teaching in classrooms are that the interaction between tutors and students often is one-to-many. In addition, language teaching does not reflect how language is naturally acquired and the tutor is often either not fluent in the second language or not proficient in the child’s mother tongue. While there is large variation in L2 proficiency in young children, with factors such as gender, socio-economic background and home education having a significant impact, there is ample evidence for the current language education provision and the young learners’ subsequent L2 performance being on occasion suboptimal (Brühwiler and Blatchford, 2011; De Feyter and Winsler, 2009; Kim et al., 2014). While a number of educational approaches

adfa, p. 1, 2011.

remedy this through, for example, immersion approaches, second language teaching remains challenging, especially for immigrant children (Collins et al., 2012).

It has long been established that one-to-one tutoring can result in significantly higher cognitive learning gains than group education. Bloom (1984) found that one-to-one tutoring resulted in 2 standard deviations improvement against a control group, concluding that "the average tutored student was above 98% of the students in the control class" (p. 4). Whilst research since has shown that the effects are not as large as first observed, there is nonetheless a distinct advantage to the one-to-one tutoring approach (VanLehn, 2011). However, traditional school classroom arrangements mean that one teacher is responsible for many children. In such situations it is not possible for teachers to offer as much one-to-one tutoring as would be desired.

More recently it has emerged that social robots can be used to teach children and adults. However, what is remarkable is that social robots seem to have a distinct advantage over alternative digital one-to-one tutoring technologies, such as screens and tablets. When tutoring is delivered by a social robot this leads to greater learning gains compared to the same content delivered on-screen (Han et al., 2005, Hyun et al., 2008, Kose-Bagci et al., 2009, Leyzberg et al., 2012), with performance increases of up to 50% compared to interactive screen technology (Kennedy et al. 2015). The reasons for this are still unclear: it might be that the social and physical presence of the robot engages the learner more than just on-screen delivery and feedback, or it might be that the learning experience is a more multimodal experience thus resulting in a richer and embodied pedagogical exchange (Mayer & DaPra, 2012), or of course a combination of these two.

Of importance here is that robots (and digital media such as tablets and computers) allow for a fast-paced interaction, and digital devices can tailor the interaction to match the level and interests of the young learner. This allows for the system to stay within Vygotsky's Zone of Proximal Development of the child and adopts an interactionist perspective to learning (Chapman, 2000); both approaches are central to this project.

L2TOR (pronounced 'el tutor'), runs for 3 years starting early 2016, and aims to design a child-friendly tutor robot that can be used to support teaching preschool children a second language by interacting with children in their social and referential world. In particular, the project will focus on teaching English as L2 to native speakers of Dutch, German and Turkish, and teaching Dutch and German as L2 to immigrant children speaking Turkish as a native language. The L2TOR robot will be designed to interact naturally with children aged four years old in both the second language and the child's native language. The robot's social behaviour will be based on how human tutors interact with children, and will not only use verbal communication, but also non-verbal communication, such as gestures and other forms of body language. The robot will be able to adaptively respond to children's actions and engage with them in tutoring interactions. The child will be provided with increasingly complex stimuli and utterances in the second language, as well as appropriate feedback that support the child's language development.

2 General approach

The central goal of the L2TOR project is to develop an embodied digital learning environment in which a child-friendly, social humanoid robot serves as a tutor to assist children acquiring a second language. This robot will be able to interact with the child naturally at a level that challenges the child to learn new words and grammar, while at the same time feels like a friend. The robot will keep track of individual children's development and will adapt its own interaction to facilitate the child to advance to the next level. As such, the robot will construct a scaffold that allows the child to acquire new skills in interaction. Since the robot will teach the child a second language, proficiency in the child's native language is desirable, so it can provide explanations and instructions that the child can readily understand.

The L2TOR embodied digital learning environment will not only consist of the robot, but it is a complete learning environment that also consists of a table-top environment that represents the contextual content of the system. Depending on the educational domain, this table-top environment will either be a table with moveable objects or an interactive tablet computer (Fig 1). Together with the child, the robot and table-top will constitute the contextual setting in which the tutoring will take place.

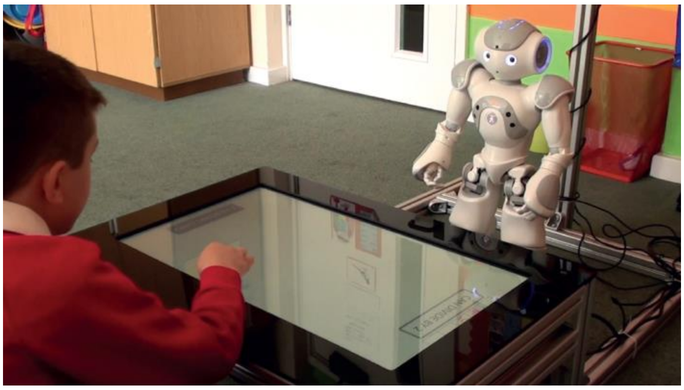

Fig. 1. A robot teaching division skills and prime numbers to a primary school pupil (Kennedy et al., 2015). L2TOR will use a similar setup, using a tablet computer instead of a larger display.

To develop an effective tutoring robot, the robot should interact with a child in similar ways a caregiver or teacher would do when teaching the child language. Such interactions not only include verbal content, but also nonverbal content and adequate socio-cognitive skills, because these form the pragmatic backbone of language acquisition from infancy on (Matthews, 2014). This multimodal interaction allows the interactants to construct and maintain common ground, which is essential for language learning, because this provides the child with a suitable context to learn from. Since there are few observational data on multimodal interaction for L2 language tutoring, we will collect our own data and analyse these such that they can be incorporated as a template for the L2TOR robot.

The primary requirement for building common ground is to design child-robot interactions that allow for mutual understanding of the communicative acts and the environment in which the interactants are situated. For the L2TOR robot, this means that the robot should be able to

- perceive and recognize the objects and events that occur in the environment,
- perceive and recognize the verbal and nonverbal signals produced by the child,
- use Theory of Mind to take the child's perspective,
- be able to monitor linguistic/behavioural errors produced by the child,
- respond to the child in a contingent manner, both temporally and semantically and,
- produce appropriate utterances in different modalities (particularly, speech and gesture) and in different languages (native and target language).

The design of these capacities is the major challenge that the L2TOR project needs to tackle, but current technology is sufficiently advanced to provide pragmatic solutions for most issues. For example, the perception and recognition of social signals is unsolved for open domains, but early work shows that for closed-domain interactions, we have sufficient interpretability to allow for full autonomy (Kennedy, Baxter & Belpaeme, 2015). As far as possible, the implementation will rely on integrating existing technologies, especially for the hardware solutions, the input recognition and the motor control of the system. A key point here will be speech recognition, with current speech recognition system not performing with sufficient reliability for child speech; to mitigate this, interaction which be directed through a touch screen interface on which the young learner taps icons. The tutor robot will be realised by Aldebaran Robotics' Nao humanoid (Fig. 1), which comes with a large range of suitable software for input and output processing. The challenges occur in the design and implementation of multimodal interactions that have the capacity construct common ground with the child to facilitate L2 acquisition.

3 Three lesson series

While interaction design for robots has been explored extensively, research into how interactions should be designed to support tutoring and teaching is recent and as of yet inconclusive. As a first goal, the pedagogy of robot assisted language tutoring will have to be defined. For this purpose, the L2TOR project will design, implement and evaluate three series of lessons (each running 10-15 weeks, 3-4 sessions per week) for the three educational domains:

1. Number domain: Learning language about basic number and pre-mathematical concepts.
2. Space domain: Learning language about basic spatial relations.
3. Storytelling domain: Vocabulary and concept learning during storytelling.

These domains were chosen to restrict the range of interactions such that the objectives are feasible and measurable within the duration of the project, while at the same time being relevant and suitable for educational purposes in a pre-school setting. Each lesson will be implemented and evaluated for five language pairs L1 and L2: native speakers of German, Dutch and Turkish will be taught English, while Turkish (immigrant) children will be taught Dutch or German, depending on their country of residence. These language combinations are not only chosen for practical considerations (they cover the languages of the academic partner states involved), but also for strategic reasons. First, English is the most commonly taught second language across Europe. Second, many children of Turkish immigrants live in the Netherlands and Germany, and will learn Dutch or German in preschool and beyond. Thus, the latter will represent a common situation of ethnic minority children learning L2 at school.

For each domain, learning targets will be developed. In the number domain, the learning targets will increase in complexity from mere counting objects and naming of shapes, to comparing numerosities, and to performing transformations on objects and sets (addition, subtraction, identifying geometrical shapes). In the space domain, learning targets range from exploring spatial relations between objects from an egocentric perspective (preposition and movement verbs), to spatial relations from an allocentric perspective (navigation through space and perspective taking), and performing a construction task (building a model with blocks) following instructions involving spatial relations, spatial coordinates and movement through space. The learning targets for the storytelling domain include vocabulary about rare objects and events (e.g., “wooden bird”, “magical flying bird”), and basic narrative structures.

For each lesson series, the L2TOR robot will communicate with the child following a specified scenario to obtain the learning targets. These scenarios describe the general sequence of targets that L2TOR aims to achieve by interacting with the child. The scenarios need to be adaptive, because the interactions between child and robot are adaptive and to some extent unpredictable. The contexts for the number and space domains are provided by a blocks/toy world that the children and -to a limited extent- the robot can manipulate. For the number domain, scenarios will be designed in which the objects can be grouped in countable sizes. For the space domain, blocks can be positioned in different ways (e.g., putting blue block on the red block) to test children’s use of spatial language for spatial relations between objects. In the storytelling domain, the L2TOR will show the child on the tablet a story about a (currently not available) “magical transformation machine”, where a character (e.g. a wooden bird) chooses an object among several objects, puts is through a device and transforms into another object (e.g. a flying animate bird). The children will first be asked to form narratives about what they have watched. Later, the child will be given the opportunity to join in a different version of this story with the characters and actions of her own choice.

4 References

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational researcher*: 4-16.

Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, 21(1):95-108.

Chapman, R. S. (2000). Children's language learning: An interactionist perspective. *Journal of Child Psychology and Psychiatry*, 41(1):33-54.

Collins, M.F. (2010). ELL preschoolers' English vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly*, 25, 84–97.

De Feyter, J. J., & Winsler, A. (2009). The early developmental competencies and school readiness of low-income, immigrant children: Influences of generation, race/ethnicity, and national origins. *Early Childhood Research Quarterly*, 24(4):411-431.

Han, J., Jo, M., Park, S., and Kim, S. (2005). The educational use of home robots for children. In *Proceedings of the 14th IEEE International Symposium on Robots and Human Interactive Communications, RO-MAN 2005*, pages 378-383. IEEE.

Hoff, E. (2013). Interpreting the Early Language Trajectories of Children from Low SES and Language Minority Homes: Implications for Closing Achievement Gaps. *Developmental Psychology*, 49(1):4–14.

Hyun, E., Kim, S., Jang, S., and Park, S. (2008). Comparative study of effects of language instruction program using intelligence robot and multimedia on linguistic ability of young children. In *Proceedings of the 17th IEEE International Symposium on Robots and Human Interactive Communications, RO-MAN 2008*, pages 187-192. IEEE.

Kennedy, J., Baxter, P., and Belpaeme, T. (2015). The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning. In *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction*, pages 67-74. ACM.

Kim, Y. K., Curby, T. W., & Winsler, A. (2014). Child, family, and school characteristics related to English proficiency development among low-income, dual language learners. *Developmental psychology*, 50(12):2600.

Kose-Bagci, H., Ferrari, E., Dautenhahn, K., Syrdal, D. S., and Nehaniv, C. L. (2009). Effects of embodiment and gestures on social interaction in drumming games with a humanoid robot. *Advanced Robotics*, 23(14):1951-1996.

Leseman, P. P. M., & van den Boom, D. C. (1999). Effects of quantity and quality of home proximal processes on Dutch, Surinamese–Dutch and Turkish–Dutch preschoolers' cognitive development. *Infant and Child Development*, 8(1):19-38.

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, CogSci 2012*, pages 1882-1887.

Matthews, D. (Ed.). (2014). Pragmatic development in first language acquisition (Vol. 10). John Benjamins Publishing Company.

Mayer, R. E., & DaPra, C. S. (2012). An Embodiment Effect in Computer-Based Learning With Animated Pedagogical Agents. *Journal of Experimental Psychology*: *Applied*, 18(3), 239–252.

VanLehn, K. (2011) The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4):197-221.