Observing human tutoring to develop robot-based language lessons

Decades of research by psychologists and educationalists identified a number of strategies human adults use to teach a new language to young children (e.g., Konishi et al., 2014). In recent years, scholars in robotics and related fields have also been involved in research on early language education, advocating the potential of humanoid robots as companions that simulate the way human adults teach language (e.g., de Haas et al., 2016). To this date, however, there has not been extensive discussion on how strategies employed by human teachers can be applied to develop robot-based language lessons. In this project, we aim to determine which pedagogical approaches can be and should be implemented in robot-based language lessons.

This abstract discusses our first step – analysis of teaching strategies observed in preschool language classes, with the special focus on the use of first (L1) and second (L2) languages and bodily actions such as gestures. We chose these two topics due to the potential strengths of a robot as a language teaching tool. First, the ability to switch between L1 and L2 makes robots as effective as or perhaps more effective than human teachers in some situations. It is often difficult for a human teacher to switch between two languages especially when the classroom consists of children from different language backgrounds. A robot can provide supplementary one-on-one L2 lessons using any L1-L2 combinations. Second, the capability to perform actions makes a robot different from other devices such as a tablet. As a physical agent with arms and legs, humanoid robots are able to perform various gestures which are, at least when performed by humans, known to facilitate language learning in children (e.g., Hostetter, 2011; Sueyoshi & Hardison, 2005).

To assess how human L2 teachers use language and actions, we conducted semi-naturalistic observations of (1) large-group L2 English lessons at preschools in Turkey, (2) one-on-one or small-group L2 English lessons in the Netherlands and Germany, and (3) L2 Dutch lessons for children from immigrant families in the Netherlands.

First, teachers' utterances were transcribed, and all bodily actions were noted alongside. We then coded each utterance for a number of characteristics using an original coding scheme. All utterances were coded for whether it was in L1 or L2, and for whether switching between L1 and L2 occurred. Gestures and other actions were classified at different levels. At a global level, gestures can be classified into categories such as a *deictic gesture* (pointing at different entities such as objects or locations), an *iconic gesture* (gesture that represents a concrete event or object), or a *metaphoric gesture* (gesture that represents an abstract concept such as knowledge). It was also useful for our purpose to note more specific categories that can be directly used in our robot-based lessons. Thus, our codes included both general categories (e.g.,deictic gestures) and specific categories (e.g., pointing to a box, pretending to wear a jacket). Some of these codes were derived from the literature whereas others were added by our coders based on the observations. We also coded non-gestural actions (e.g., dancing on a song) because most commercially available humanoids (e.g., Softbank Robotics NAO) are expected be able to perform them.

Our observations show that, in terms of language use, English teachers in Turkey and the Netherlands mainly used the L2 as the medium of instruction. However, the teachers sometimes shifted from L2 to L1 (1) to manage classroom issues, (2) to ask questions, (3) to give instructions, and (4) to explain syntactic or phonological rules (e.g., explaining the difference between 'this is' vs. 'these are' or explaining 'the singular-plural distinction' as in shoe vs. shoes). However, in Germany, the teacher switched very frequently between the L1 and L2: out of all utterances: 55% was in L1, 30% was in L2, and 15% was unclassifiable (e.g., interjection, children's names). We can claim that teachers were naturally adjusting

their language in order to ensure that children understood key concepts. Although many L2 programs take, or at least claim to take, a total immersion approach in which the teacher speaks only in L2, the use of L1 can be still observed and is believed to be quite beneficial in some situations (e.g., Moore, 2002).

The use of gestures and other actions was very frequent in all lessons. The amount, however, varied greatly across lessons, from 9.24 to 73.07 per 20 minutes. Importantly, the rate of action use seemed to depend largely on the theme of a lesson. For instance, when teaching names of body parts, 70% of the teacher's utterances containing target words were accompanied with gestures (e.g., pointing to the arms), but when the lesson theme was weather, the teacher used gestures in only 18% of her target word utterances. Thus, teachers used gestures only when there was a conventional or very straightforward gesture associated with the word they were teaching.

Our data suggest that, although the mere presence of gestures may increase children's attention to the learning content (e.g., Hostetter, 2011), including gestures in every possible occasion may not be necessary. Gestures are suited to teach words in some domains such as math (Cook & Goldin-Meadow, 2006), but may not be as important when teaching concepts with no conventional gestures such as body parts or color because the behavior was not observed among human teachers either. It must be noted, however, in teaching any concepts, some gestures such as pointing can be useful. Pointing can direct children's attention to any relevant object, material, or location. In fact, pointing was more common than any other gestures (e.g., iconic gestures) in our class observation.

So how can we use the information in L2 classrooms to develop a robot L2 tutor? Our data on language use suggests that L1 is used in L2 classrooms more often than commonly believed, and the amount of L1 use is flexibly determined based on various factors such as lesson topics and L2 proficiency of students. The results also highlight the potential benefits of using a robot as a language tutor because, as mentioned earlier, a robot can be programmed to use any combination of L1 and L2 in theory.

Translation of the pedagogical strategies used by human teachers to robot-based lessons also introduce unique challenges. Although we found that the teachers constantly performed actions to facilitate their learning process, the robot gesturing too much might cause more harm than good. Most humanoids available under the status quo cannot move as flexibly or smoothly as humans, and thus some of the gestures observed in the classrooms cannot be well replicated by robot tutors. Further, many robots produce motor sounds while gesturing and thus can mask speech sound when utterance and gesture simultaneously occur. Research suggests that overuse of actions and gestures or mismatch between speech and gesture can impede the word learning process (e.g., Goldin-Meadow & Sandhofer, 1999). Even human teachers do not perform actions in some situations, and thus in designing robot-assisted L2 lessons, we must carefully consider when the use of actions and gestures is truly appropriate, as opposed to including them as much as possible.

In conclusion, we emphasize that observation of human tutoring can be quite beneficial in developing robot learning companions not only because it provides general ideas about how children learn a new language, but also because specific phrases and actions used by human teachers are most likely to be familiar for children and thus may help children recognize the robot tutor as an agent and to have a successful learning experience. With regards to some features such as gestures, we must carefully consider the balance between what we want the robot to do and what hardware and software limitations of the particular robot let us do. Observations of human tutoring can serve as a good starting point in determining what to consider in developing educational robots.