Second Language Tutoring using Social Robotss

Project No. 688014

L2TOR

Second Language Tutoring using Social Robots

Grant Agreement Type: Collaborative Project
Grant Agreement Number: 688014

D1.2 Semi-naturalistic observations of human tutoring

Due Date: **28/02/2017**
Submission Date: **14/03/2017**

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

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

Responsible Person: **Aylin Küntay**   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

The aim of this deliverable is to discuss the characteristics of teacher-child interactions in preschool L2 lessons. The primary focus was given to English lessons in Turkey, but we also observed children learning English in the Netherlands and Germany as well as immigrant children learning Dutch in the Netherlands. This document describes how the observations were conducted and reports a variety of factors including use of L1, actions and gestures, and attention grabbers. The observations will give us insight into the way L2 tutoring is conducted as well as into how child-robot interactions may be implemented in these settings.

Principal Contributors

Koç   Aylin Küntay, Tilbe Göksun, Junko Kanero, Vasfiye Geckin, Cansu Oranç, Ezgi Mamus

Utrecht   Paul Leseman, Rianne Vlaar, Ora Oudgenoeg-Paz, Josje Verhagen

Tilburg   Paul Vogt, Emiel Krahmer, Mirjam De Haas

Bielefeld   Stefan Kopp, Kirsten Bergmann, Thorsten Schodde

Plymouth   Tony Belpaeme

Revision History

Version 1.0 (AK 14-03-2017)
First version.

1 Introduction

   With the increase in the number of immigration and global domination of English as an academic language, teaching second language (L2) to young children is of utmost importance. Many preschools in Europe and elsewhere have integrated L2 English into their curricula; however, it is impossible for preschools to provide language input that is quantitatively and qualitatively sufficient for successful L2 acquisition. To supplement L2 learning at school, digital media tools including educational social robots are being developed (e.g., Kanda et al., 2012). Importantly, benefits of using digital devices in early education cannot be maximized unless lessons are engaging and enjoyable and thus appropriate for young learners. Observations of human tutoring, especially of experienced and skilled teachers, can provide useful insights to how a social robot can be useful for language tutoring. We conducted semi-naturalistic classroom observations in order to assess how human teachers teach L2 to young children (The prefix "semi-" is used here because the teachers were informed that they were being filmed). This deliverable reports data from these observations – both verbal and nonverbal interaction patterns between teacher and child – that can be used in designing our robot-assisted lessons.

   All observations reported here were coded according to the coding scheme developed for this project. This coding scheme is designed to assess both verbal and nonverbal characteristics of human L2 tutoring, including various types of feedback and teachers' use of gestures (see Appendix 1 for the full coding scheme). Previous research suggests that teachers use various types of feedback to facilitate L2 learning and to boost motivation of children (e.g., Noor, Aman, Mustaffa & Seong, 2012). For instance, when a learner says "She play tennis," the teacher may correct this ill-formed sentence by saying "It's not: She play, but: She plays." Alternatively, the teacher may add metalinguistic comments. For example, when a student says "Last night rained," the teacher may explicitly correct the child by saying "We use 'it' when talking about the weather." To help students learn the L2, teachers may also use implicit feedback. For example, repeating the student's utterance with prosodic marking on the erroneous part of the sentence (e.g., "I BROKED a vase?") could lead the learner to self-correct his or her utterance. Finally, the teacher may use facial expressions like frowning and other body movements together with or instead of verbal feedback (Yao, 2000).

   Gestures and actions are also important for L2 tutoring, especially for the L2TOR project because the ability to gesture and perform actions is what makes social robots unique and different from other digital devices such as tablets. Human teachers use non-verbal cues such as gestures as pedagogical tools to make input more comprehensible to the learners (e.g., Taleghani-Nikazm, 2008), to expand learners' verbal expressions (Quinn Allen, 1995), to help learners self-correct their errors (Muramoto, 1999), and to create a positive learning environment lending room for scaffolding (McCafferty, 2002). Previous research indicates that gestures play an important role in L2 learning (see Gullberg, 2008 for a review). For example, 5-year-old French-speaking children learning English recalled more English words after watching and producing the accompanying gestures than just watching the gestures without speech (Tellier, 2008).

   In the project proposal, we originally proposed to video tape and analyse human tutoring in two settings: 10 hours of L2 Dutch and German lesson for ethnic minority children of Turkish origin in the Netherlands and Germany and 10 hours of English lesson for children in Turkey, the Netherlands, and Germany. This plan, however, was

modified because preliminary data analysis based on 6 hours of lesson observation suggested that teaching strategies did not differ drastically across teachers or lessons, and thus 20 hours of observation was not necessary. We decided to reduce observation hours but analyse each lesson in depth. As our plan for the whole project is to first develop English lessons and then move onto non-English lessons, in Task 1.2 reported here, we focused primarily on how human teachers taught English to young non-English speakers. To take an in-depth look at each lesson, video samples from English lessons were annotated not only in terms of vocabulary, multimodal interaction, the use of L2, and reaction of the learners as planned originally, but also in terms of other factors such as feedback and attention grabbers.

For all analyses, English lessons in Turkey were analysed primarily as those videos were obtained specifically for this project and thus complete information was available for these teachers and lessons. We additionally analysed existing video materials of English lessons in the Netherlands and Germany to evaluate the generalizability of the patterns found in the Turkish classrooms. L2 Dutch lessons for immigrant children were also analysed as this project plans to design L2 Dutch and German lessons later in the process. However, as non-English lessons are not our primary focus at this moment, we report six L2 Dutch sessions only at this moment.

This deliverable consists of three main sections. First, we provide a general overview of how the observations were conducted. Second, we report the results of a quantitative analysis of the data obtained. The quantitative analysis includes factors such as the frequency with which specific, targeted L2 words were taught, hereafter *target words*, and the frequency of L1 use in classrooms. The third part of this deliverable presents the results of a descriptive analysis focusing on qualitative components of human L2 tutoring that cannot be fully captured with a quantitative analysis.

2 Overview of the Semi-Naturalistic Observation

2.1 Filming of the Lessons

The observational studies were conducted in Turkey, Germany, and the Netherlands. For L2 English lessons in Turkey, two preschools in Istanbul participated in the study. We first obtained ethics permission and consents from the schools, teachers, and parents. The teachers were informed on the observation beforehand and were aware that they would be filmed. Cameras were placed in the corners of the classroom to minimise the distraction caused to the children. For the other settings (L2 English lessons in the Netherlands and Germany and L2 Dutch lessons in the Netherlands), we analysed video clips that were obtained in previous studies. The teacher utterances in all of the films were transcribed and coded.

2.2 Lesson Overview

Table 1 shows an overview of all lessons observed in this task. In Turkey, researchers from Koç University (KOC) visited two preschools and observed three lessons in each school. The same teacher taught all three lessons at each school (Teacher 1 and Teacher 2). In the Netherlands, three video samples of English lessons were obtained from previous research projects conducted at Utrecht University (UU) and Radboud University. All lessons were taught by different teachers. Six video samples of L2 Dutch lessons were also obtained from a previous project at UU which focused on effects of preschool education in mixed and targeted classroom (see de Haan,

2015 for the details of the project). All lessons were taught by different teachers, except one teacher (Teacher 9) who was observed in two sessions. Finally, an English lesson in Germany was acquired from an L2 tutoring session observed by the University of Bielefeld. Table 1 lists the IDs of the teachers, the L2 taught in lessons, the countries in which recordings were obtained (TR = Turkey, NL = the Netherlands, DE = Germany), the duration of the lessons, and the target words taught in the lesson. As stated above, our analyses primarily focus on L2 English lessons in Turkey. But, transcripts of all other lessons were also analysed to examine the generalizability of our findings.

Table 1. The overview of all lessons observed in Task 1.2.

Teacher ID	L2	Country	Lesson duration (min)	Target words taught
Teacher 1	English	TR	41:54	head, shoulders, knees, toes, eyes, ears, mouth, nose, face, finger, chin, cheek, arm, leg, foot, hair, hand
Teacher 1	English	TR	38:56	hat, scarf, shoes, jacket, gloves
Teacher 1	English	TR	35:40	yellow, brown, green, red, blue, white, orange, black, purple, pink
Teacher 2	English	TR	57:30	duck, horse, sheep, cow, rabbit, pig, tiger, jungle, cat, dog, bird, monkey, zebra, giraffe, crocodile, farm, snake, spider, elephant, lion, turtle, house, home
Teacher 2	English	TR	36:48	polar bear, penguin, sheep, horse, snake, fish, bird, zebra, duck, lion, pig, rabbit, cat, giraffe, tiger, elephant, spider, crocodile, pet, animal, jungle, farm, lake, house
Teacher 2	English	TR	28:07	polar bear, penguin, sheep, horse, snake, fish, bird, zebra, duck, lion, pig, rabbit, cat, giraffe, tiger, elephant, spider, crocodile, pet, animal, jungle, farm, lake, house, crab, shark, octopus, jelly fish, waves, lobster, squid, whale, dolphin, seal, turtle, walrus, ocean, monkey, shrimp, starfish
Teacher 3	English	NL	19:03	pink, purple, black, yellow, white, red, circle, brown, blue, heart, square, star, rectangle, moon, oval, diamond, triangle
Teacher 4	English	NL	19:08	Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday, sun/sunny, clouds, dinner, breakfast, lunch, food, pasta, cake, rice, salad, cheese, sandwich
Teacher 5	English	NL	28:04	weather, cold, hot, windy, snowy, cloudy, rainy, misty, sunny, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday, umbrella, lemon, strawberry, orange, apple, pear, banana, one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen
Teacher 6	English	DE	27:57	eye, legs, feet, toes, knees, ears, nose, hands, fingers, lips, teeth, tongue, head, hair, where are your...?

Teacher 7	Dutch	NL	10:12	Words relating to the Sinterklaas holiday: Sinterklaas, Black Pete [Zwarte Piet], Beard [baard], Horse [paard], Steamboat [stoomboot], Moustache [snor], Present [kadootje], Cane [staf], roe (no translation available, it is a bundle of twigs carried around by Black Pete), book [boek], Amerigo (name of a horse),collar [kraag], shoe [schoen], hat [muts], feather [veer], bag [zak]
Teacher 8	Dutch	NL	25:51	Pizza [pizza], king [koning], queen [koningin], restaurant [restaurant], table [tafel], telephone [telefoon], cloak [mantel], crown [kroon], waiter [ober], chef [kok], towel [theedoek], table cloth [tafelkleed], plate [bord], dishware [servies], glass [glas], fork [vork], knife [mes], spoon [lepel], to lay the table [tafel dekken]
Teacher 9	Dutch	NL	22:10	Weather [weer], sun [zon], rain [regen], clouds [wolken], snow [sneeuw], wind [wind], lightning [bliksem], hot [warm], cold [koud], summer [zomer], autumn [herfst], winter [winter], leaves [blaadjes], rake [hark], ladder [ladder], light [licht], flower [bloem], pile [hoop/berg], up [omhoog], down [omlaag[
Teacher 9	Dutch	NL	23:46	Puppy [puppy], weather [weer], lightning [bliksem], thunder [onweer], wind [wind], (white/grey) clouds [(witte/grijze) wolken], sun [zon], snow [sneeuw], hail [hagel], snowman [sneeuwpop], hat [muts], shawl [sjaal], gloves [handschoen], coat [jas], nice weather [lekker/mooi weer], hot [warm], cold [koud], summer [zomer], autumn [herfst], winter [winter], spring [lente], tree [boom], to shine [schijnen], leaf [blaadje], to blow [blazen], up [omhoog], roof [dak]
Teacher 10	Dutch	NL	14:30	Puddles [plassen], raining [regenen], boots [laarzen], coat [jas], wellies [regenlaarzen], rain coat [regenjas], umbrella [paraplu], scooter [step], mud [modder], to wash [wassen], bath [bad], clothes [kleren], soap [zeep], cold [koud], hot [warm], to shower [douchen], to undress [uitkleden], spider webs [spinnenwebben], spider [spin], cow [koe], sheep [schaap], goat [geit], pig [varken], duck [eend], cock [haan], dry [droog]
Teacher 11	Dutch	NL	7:45	Bear [beer], stone [steen], foot [voet], head [hoofd], ambulance [ambulance], ziekenwagen (other word for ambulance), white [wit], red [rood], hospital [ziekenhuis], bandage [verband], tears [traantjes], present [cadeautje], doctor's coat [doktersjas], stethoscope [stethoscoop]

Note: TR = Turkey, NL = the Netherlands, DE = Germany

2.3 Teacher-Child Profiles

Six observations in Turkey were conducted in two different preschools with children between the ages of 60 and 72 months. Each teaching session lasted around 40 minutes. The classrooms had a vivid atmosphere with lots of posters and pictures on the wall. Data came from two Turkish-speaking instructors who had a BA degree in 'Teaching English as a Foreign Language' (TEFL) from a Turkish university. Both teachers were second language speakers of English (see Table 2 for details). Teachers

designed their own syllabi and activities. In each school, children were exposed to L2 for about 40 minutes a day, and there were sometimes two 40-minuute L2 sessions in one day. We do not have this information about the teachers in the Netherlands as these observations come from previous projects which did not document these details.

Table 2. Profile of teachers and students.

Session #	Class size	School	Teacher	Teaching experience
1	9 (3 female)	School 1	Teacher 1	15 years
2	6 (1 female)			
3	9 (2 female)			
4	12 (6 female)	School 2	Teacher 2	5 years
5	16 (9 female)			
6	12 (8 female)			
7	27 (19 female)	NA	Teacher 3	NA
8	20 (6 female)	NA	Teacher 4	NA
9	26 (13 female)	NA	Teacher 5	NA
10	5 (2 female)	NA	Teacher 6	NA
11	4 (3 female)	NA	Teacher 7	NA
12	4 (2 female)	NA	Teacher 8	NA
13	5 (2 female)	NA	Teacher 9	NA
14	3 (1 female)			
15	2 (1 female)	NA	Teacher 10	NA
16	4*	NA	Teacher 11	NA

*Gender composition of Session 16 is unknown.

2.4 Development of the Coding Scheme

The coding scheme was developed by KOC based on the observations conducted in two preschools in Turkey as mentioned above. There are existing coding schemes examining teaching strategies; however, we did not find a coding scheme that fits the L2TOR project as our aim is to design a robot based on the obtained data. Therefore, we decided to develop our own coding scheme that was based on previous literature as well as our own data. First, teachers' utterances were transcribed. Teachers' speech was transcribed in utterances – small chains of words, phrases, or sentences – based on interruption in speech sound, change in pitch, and speech content. Then, we scored each utterance for various characteristics such as the use of actions and gesture, types of feedback, and attention grabbers and noted all the categories we observed. Some of these codes in the scheme were derived from the literature (e.g., pointing, negative feedback) whereas others were added by our coders based on the observations (e.g., showing a flashcard). Once these characteristics were noted, we further labelled patterns that appeared consistently across the sessions. To evaluate the reliability of the coding scheme, at least 10% of the observational data was coded by multiple coders, and any mismatches between the coders was discussed and modifications were made in the coding scheme. Once the reliability among the coders in Turkey passed the threshold of 80%, we begun using the coding scheme in the Netherlands and Germany for further refinement. The final version of our coding scheme consisted of five main categories described below (see Appendix 1 for the full coding scheme). The final coding scheme was used to code all observations in all three countries.

2.4.1 Language use

The first category of the coding scheme was language use. A major decision educators and curriculum developers must make in designing lessons is deciding how much L1 and L2 should be used in the lesson. It is good to maximize the amount of L2 exposure by using only L2. However, especially with children who just started to learn an L2, hearing everything in L2 can be confusing and frustrating. To examine how human teachers deal with the issue of L1 vs. L2, we examined the amount of L1 and L2 used by each individual teacher. Teacher utterances in the second language were coded as "L2" and teacher utterances in the children's native language were coded as "L1." Within certain utterances, the teacher used both L1 and L2, i.e., code switches. To understand the nature of code switching, we marked *intrasentential code switching* in which the teacher alternates languages within a sentence, and *intersentential code switching* in which the teacher alternates languages across sentences. The frequency of L1 and L2 use as well as code switches are reported here. Note that there was no code switching in L2 Dutch lessons because the teacher did not speak the children's L1.

2.4.2 Feedback patterns

As discussed briefly in the introduction, effective feedback is essential for L2 learning. The feedback patterns observed were coded as two categories: *positive feedback* and *negative feedback*. Positive feedback constitutes words, sentences, or phrases that affirm children's speech and behaviors. Affective feedback in the form of praise can be considered as positive feedback (e.g., Well done! Good job! Awesome!), encouragement (e.g., That's it! You can do it!), confirmation (e.g., That's right!) or consolidation (the teacher repeats the correct response of the child; e.g., Yes, the cat is in the house!). These forms of feedback are used to motivate the child to continue the task. Negative feedback, on the other hand, is provided when the child's behavior is undesirable, mostly when she gives a wrong answer to a question. Negative feedback can be simply saying "no." But, the teacher can provide more informative corrective feedback, either implicitly or explicitly. Thus, negative feedback was further divided into these two categories. In explicit corrective feedback, the teacher overtly and clearly indicates that there is a specific error in the child's response (e.g., No, this is wrong, it is not a cat, it is a dog). Implicit corrective feedback, however, usually includes "recasts." The teacher may reformulate all or part of a child's utterance (*expansion*), ask the child to repeat or reformulate her statement (*clarification requests*), or let the child discover her own mistake by repeating the child's utterance with the ill-formed part emphasized (*repetition*).

2.4.3 Actions and gestures

Gestures can be classified into various categories, and many different gestures can be used to refer to just one word. An *iconic gesture* describes a concrete event or object while a *metaphoric gesture* depicts an image of an abstract concept such as knowledge. A *deictic gesture* refers to pointing at different entities, objects or locations (McNeill, 1992). Deictic gestures are frequently used to inform and assess the students and to organize the lesson (Azoui, 2013). For these different gestures, we noted a general category (e.g., pointing) and a more specific description that specifies the referent of the gesture (e.g., pointing to the box). We coded not only gestures (e.g., pointing, pretend play) but also other actions (e.g., dancing along with the lyrics of a song) as our robot – the Softbank Robotics Nao robot – can perform non-gestural

actions as well. Actions and gestures were marked for both L2 target words and other L2 utterances (Appendix 1).

2.4.4 Teacher-child interaction patterns

The sessions observed in the L2 lessons in Turkey had a clear opening and closure section and transitions between activities (and this was mostly the case for the lessons in the Netherlands and Germany as well). The children were aware of signals starting and concluding a task. The teachers always stated the objective of the lesson — e.g., "Today, we will learn [sic] our clothes, winter clothes." Games, songs, colouring, cutting and pasting activities constituted the practice section of the lesson. The closure part usually included saying goodbye and tidying up the classroom for the next lesson. There were also very few utterances of teachers explicitly directing children to communicate in L2. One teacher simply ignored all the questions the children asked in L1. The most frequently used interaction patterns by the teachers were giving/repeating instructions, chanting and repetition, and asking questions.

2.4.5 Attention grabbers

Attention grabbers are used to draw the attention of children when they are not paying attention and can be both verbal and nonverbal. Verbal attention grabbers are words or phrases such as "look" and "listen." Non-verbal attention grabbers are actions and gestures such as clapping hands. We marked all attention grabbers, and in addition, we rated the effectiveness of attention grabbers on a scale from 0 to 2 (0 = ineffective, 1 = somewhat effective, 2 = effective). The coders judged the effectiveness of attention grabbers based on children's reactions. We included the effectiveness ratings in the coding scheme. Although the ratings were highly subjective and could not be used to statistically assess the effectiveness of a particular attention grabber, they could still provide an idea about which attention grabbers might be more useful than others.

3 Quantitative Analysis

In this section, we summarize the results of the quantitative analysis that evaluated the frequency of use of the various teaching strategies in L2 classrooms. As the duration of the lessons varied across sessions, we calculated how often each strategy was used per utterance or minute. Language use (how much L1 and L2 were used in lesson) was estimated by dividing the number of utterances in L1 and, respectively, L2 during a lesson by the total number of utterances in a lesson (Table 3). For the use of actions and gestures, we estimated how often (1) utterances containing target words (i.e., target-word utterances), (2) utterances containing any L2, and (3) L1 utterances were accompanied by gestures and actions (Table 6). For example, the rate of action and gesture use for target-word utterances was calculated by dividing [the number of target-word utterances accompanied by actions/gestures] by [the number of all target word utterances]. For all other coding categories, we present numbers in a "per 20 minutes" format. We chose 20 minutes as a reference duration because, in the L2TOR project, we are planning to design a series of 20-minute robot-assisted L2 lessons. Those numbers indicate the expected frequency of each teaching strategy if the teachers were to teach a 20-minute lesson. For all coding categories, we also calculated proportion of target utterances (the number of utterances containing a word or phrase of interest such as target words, feedback phrases, and attention grabbers divided by the total number of

utterances) just as we did for the language use measure. We, however, decided to report the "per 20 minutes" numbers in this deliverable because the duration of lesson can be more reliably estimated than the total number of utterances and the "per 20 minutes" numbers can be directly used to determine how we should implement what we observed as a reference when designing our robot-assisted lessons.

3.1 Language Use

Table 3 provides information about the language use in teachers' utterances per lesson and Table 4 shows the frequency of code switching. The two teachers in the Turkish preschools mainly used the L2, English, as the medium of instruction (see Table 3). The number of intra-sentential code-switches (i.e., mixing of LI and L2 within one sentence), was very limited (Table 4). The number of inter-sentential code-switches were not frequent either, but the teachers sometimes shifted from L2 to L1 (1) to resolve classroom management issues, (2) to ask questions, (3) to give instructions, and (4) to explain a syntactic or phonological rule (e.g. explaining the difference between 'this is' vs. 'these are' or explaining 'the singular plural distinction' as in shoe vs. shoes).

Table 3. Language use in teachers' utterances. The numbers presented here indicate the percentage of utterances delivered in L1 or L2.

Teacher ID	Session #	L2	Theme	L1 use (%)	L2 use (%)	Other (%)
Teacher 1	1	English	Body parts	23.87	70.47	5.66
Teacher 1	2	English	Clothes	11.15	84.00	4.85
Teacher 1	3	English	Colours	14.08	82.55	3.37
Teacher 2	1	English	Animals	0.00	100.00	0.00
Teacher 2	2	English	Animals	1.54	91.41	7.05
Teacher 2	3	English	Animals	0.29	99.56	0.15
Teacher 3	1	English	No specific theme	0.00	100.00	0.00
Teacher 4	1	English	Weather	0.00	100.00	0.00
Teacher 5	1	English	Weather	0.00	100.00	0.00
Teacher 6	1	English	Body parts	53.20	31.28	15.52
Teacher 7	1	Dutch	Holiday	0.00	100.00	0.00
Teacher 8	1	Dutch	Restaurant	0.00	100.00	0.00
Teacher 9	1	Dutch	Seasons	0.00	100.00	0.00
Teacher 9	2	Dutch	Seasons	0.00	100.00	0.00
Teacher 10	1	Dutch	Weather	0.00	100.00	0.00
Teacher 11	1	Dutch	Weather	0.00	100.00	0.00

Note: The numbers in the Other column reflect utterances that cannot be classified as neither L1 or L2. These utterances contain code-switching, interjection, and proper names (e.g., a child's name).

Table 4. The frequency of intra- and inter-sentential code-switches. The numbers in the Intra-sentential and Inter-sentential columns indicate how many times each type of code-switches would be used in a 20-minute lesson, calculated based on the rate of code-switching found in each session.

Teacher ID	Session #	L2	Theme	Intra-sentential	Inter-sentential
Teacher 1	1	English	Body parts	1.43	62.05
Teacher 1	2	English	Clothes	1.03	41.61
Teacher 1	3	English	Colours	1.68	45.98
Teacher 2	1	English	Animals	0.00	0.00
Teacher 2	2	English	Animals	1.09	3.80
Teacher 2	3	English	Animals	0.00	0.71
Teacher 3	1	English	No specific theme	0.00	0.00
Teacher 4	1	English	Weather	0.00	0.00
Teacher 5	1	English	Weather	0.00	0.00
Teacher 6	1	English	Body parts	12.10	31.96
Teacher 7	1	Dutch	Holiday	0.00	0.00
Teacher 8	1	Dutch	Restaurant	0.00	0.00
Teacher 9	1	Dutch	Seasons	0.00	0.00
Teacher 9	2	Dutch	Seasons	0.00	0.00
Teacher 10	1	Dutch	Weather	0.00	0.00
Teacher 11	1	Dutch	Weather	0.00	0.00

3.2 Target Vocabulary

Table 5 presents the frequency of theme-related target words used in a session. L2 lessons for young learners often have very specific themes such as "my body" or "my clothes," and thus we focused on those target words instead of all L2 utterances that included many unplanned instructional commands and attention grabbers. The theme-related target vocabulary items were constantly repeated in songs, chants and games. The use of songs was quite common across lessons and was used in the other classroom in Turkey as well as in lessons in the Netherlands and Germany. Some target words were repeated more frequently than others for a practical reason (e.g., the teacher sang "Head, Shoulders, Knees and Toes" and thus those four body parts were mentioned more often than names of other body parts). In addition, the teachers also repeated words that children did not seem to have memorized yet.

Table 5. The utterances containing target words. This table shows the number of target words that is mentioned by the teacher more than once.

Teacher ID	Session #	L2	Theme	# of target words	# of target words used per 20 minutes
Teacher 1	1	English	Body parts	18	13.39
Teacher 1	2	English	Clothes	6	30.48
Teacher 1	3	English	Colours	11	20.49
Teacher 2	1	English	Animals	23	4.23
Teacher 2	2	English	Animals	24	4.66

Teacher 2	3	English	Animals	27	3.14
Teacher 3	1	English	No specific theme	17	6.92
Teacher 4	1	English	Weather	20	6.53
Teacher 5	1	English	Weather	33	3.07
Teacher 6	1	English	Body parts	15	5.68
Teacher 7	1	Dutch	Holiday	16	9.68
Teacher 8	1	Dutch	Restaurant	20	4.29
Teacher 9	1	Dutch	Seasons	20	4.38
Teacher 9	2	Dutch	Seasons	28	6.67
Teacher 10	1	Dutch	Weather	26	5.73
Teacher 11	1	Dutch	Weather	14	7.93

3.3 Actions and Gestures

The use of actions and gestures was very frequent in all the lessons (Note: Some teachers did not use any L1 at all, and for those teachers, all the gestures concern L2 utterances). The amount, however, varied greatly across lessons, from 9.24 per 20 minutes to 73.07 per 20 minutes. As discussed further in the next section (4. Descriptive and Qualitative Analysis), the rate of action and gesture use seemed to depend largely on the theme of a lesson. In designing robot-assisted lessons, we need to carefully consider when the use of actions and gestures is truly appropriate, as opposed to including them as much as possible. Overuse of actions and gestures or mismatch between speech and gesture can impede the word learning process (e.g., Goldin-Meadow & Sandhofer, 1999; see Questions 8 and 9 of the next section for the descriptive analysis and Appendix 4 for examples of gestures used in lessons).

Table 6. The use of actions and gestures. The numbers in the Target words, All L2, and All L1 columns indicate how often utterances containing target words (i.e., target-word utterances), utterances containing any L2, and L1 utterances were accompanied by gestures and actions. For example, the rate of action and gesture use for target-word utterances was calculated by dividing [the number of target-word utterances accompanied by actions/gestures] by [the number of all target word utterances].

Teacher ID	Session #	L2	Theme	Target words	All L2	All L1
Teacher 1	1	English	Body parts	73.07	47.19	0.00
Teacher 1	2	English	Clothes	46.35	23.61	1.64
Teacher 1	3	English	Colours	16.67	17.35	2.56
Teacher 2	1	English	Animals	23.93	19.82	0.00
Teacher 2	2	English	Animals	7.28	15.85	16.67
Teacher 2	3	English	Animals	9.24	27.77	0.00
Teacher 3	1	English	No specific theme	26.79	32.57	0.00
Teacher 4	1	English	Weather	18.40	23.36	0.00
Teacher 5	1	English	Weather	24.65	23.87	0.00
Teacher 6	1	English	Body parts	71.43	74.45	3.00
Teacher 7	1	Dutch	Holiday	35.44	48.47	0.00
Teacher 8	1	Dutch	Restaurant	21.62	39.44	0.00

Teacher 9	1	Dutch	Seasons	72.16	36.48	0.00
Teacher 9	2	Dutch	Seasons	40.54	41.18	0.00
Teacher 10	1	Dutch	Weather	18.52	29.86	0.00
Teacher 11	1	Dutch	Weather	18.60	17.29	0.00

3.4 Attention Grabbers

Table 7 presents the frequency of use of attention grabbers. Both verbal and non-verbal attention grabbers were used very frequently in all sessions. Importantly, the lessons we observed had multiple students and thus the teachers spent a fair amount of time on classroom management. Except for the lessons of Teachers 10 and 11, these teachers interacted with only two children and as table 7 shows, had to use less attention grabbers. Because the robot is going to interact with the child one on one, while the phrases and actions used by the teachers are informative for our lesson design, we assume that the interaction between the robot and child should not require this much use of attention grabbers. The need of attention grabbers is expected vary greatly across children in our lessons, and thus we plan to estimate children's engagement by analysing children's emotion and posture (WP4).

Table 7. The use of attention grabbers. The numbers in the Nonverbal and Verbal columns indicate how many times each tWype of attention grabbers would be used in a 20-minute lesson, calculated based on the rate of attention grabber use found in each session.

Teacher ID	Session #	L2	Theme	Nonverbal	Verbal
Teacher 1	1	English	Body parts	7.16	25.78
Teacher 1	2	English	Clothes	30.82	19.52
Teacher 1	3	English	Colours	25.79	34.77
Teacher 2	1	English	Animals	11.83	45.57
Teacher 2	2	English	Animals	29.35	55.43
Teacher 2	3	English	Animals	24.90	44.10
Teacher 3	1	English	No specific theme	6.30	32.55
Teacher 4	1	English	Weather	2.09	8.36
Teacher 5	1	English	Weather	6.41	58.43
Teacher 6	1	English	Body parts	10.73	0.72
Teacher 7	1	Dutch	Holiday	17.65	3.92
Teacher 8	1	Dutch	Restaurant	21.66	3.87
Teacher 9	1	Dutch	Seasons	15.34	35.19
Teacher 9	2	Dutch	Seasons	1.68	41.23
Teacher 10	1	Dutch	Weather	2.76	2.76
Teacher 11	1	Dutch	Weather	2.58	0.00

3.5 Feedback Patterns

Table 8 presents both the feedback patterns utilized by the L2 teachers. The most frequently used corrective feedback strategy used was negative feedback[1].

Table 8. Feedback patterns in L2 teacher talk. The numbers in the Positive feedback, Negative explicit feedback, and Negative implicit feedback columns indicate how many times each type of feedback would be used in a 20-minute lesson, calculated based on the rate of feedback use found in each session.

Teacher ID	Session #	L2	Theme	Positive feedback	Negative explicit feedback	Negative implicit feedback
Teacher 1	1	English	Body parts	30.55	0.00	0.00
Teacher 1	2	English	Clothes	34.93	2.05	0.00
Teacher 1	3	English	Colours	47.66	0.56	0.56
Teacher 2	1	English	Animals	24.00	2.09	1.39
Teacher 2	2	English	Animals	36.96	2.72	8.15
Teacher 2	3	English	Animals	34.85	4.98	7.82
Teacher 3	1	English	No specific theme	107.09	1.05	11.55
Teacher 4	1	English	Weather	56.45	3.14	1.05
Teacher 5	1	English	Weather	55.58	2.14	2.85
Teacher 6	1	English	Body parts	29.34	0.00	4.29
Teacher 7	1	Dutch	Holiday	33.33	7.84	15.69
Teacher 8	1	Dutch	Restaurant	31.72	3.09	6.19
Teacher 9	1	Dutch	Seasons	82.11	1.80	0.00
Teacher 9	2	Dutch	Seasons	108.56	0.00	2.52
Teacher 10	1	Dutch	Weather	62.07	2.76	19.31
Teacher 11	1	Dutch	Weather	85.16	2.58	30.97

4 Descriptive and Qualitative Analysis

This section discusses our descriptive analysis of the classroom observations. The aim of this section is to report all qualitative components of L2 lesson that cannot be fully captured in the above quantitative analyses. This section is organized in a question-answer format so that technical developers, including our own and others, can simply look up questions that are relevant to them. Although the analyses provided in this section differ from the quantitative analyses in the previous section, we refer to the numbers presented in the quantitative section when discussing the questions. We also refer to literature whenever applicable. In the following section, we first present a list of all 13 questions we prepared, and then answer these questions one by one.

---

[1] The effectiveness of feedback in the form of deictic and iconic gestures on L2 learning in a testing session is to be reported in Deliverable 1.3.

4.1 List of Questions

1. How do teachers introduce a new activity?
2. How do teachers introduce target words for the first time?
3. What do teachers do when they need children to repeat or clarify themselves?
4. What do teachers do when children say things that are not related to the lesson?
5. What do teachers do to get children engaged?
6. What do teachers do when children are not paying attention?
7. What do teachers do when children are not answering their question? How long do teachers wait until the child responds?
8. When do teachers use gestures and how?
9. When do teachers use L1?
10. When and how do teachers use positive feedback?
11. When do teachers use negative corrective feedback to correct children's answers?
12. Why do teachers repeat some words more than other words?

4.2 Questions and Answers

1. How do teachers introduce a new activity?

There are many ways to introduce a new activity, or a task, to children. One straightforward method is to verbally describe all rules before starting a game (Teacher 2). For example, in one of the activities, Teacher 2 said "I will start the music, when I stop the music, if you have the treasure chest, you get a flash card." Teacher 7 first explained the rules of a bingo game, before actually starting the game. Another strategy was to give minimal instructions in the beginning, and letting children learn the rules as they play the game (Teacher 1). Although it is difficult to judge which method was more effective as many factors (e.g., number of children, theme of the lesson) were not controlled, the former strategy enables children to be engaged in the activity in a shorter amount of time compared to the latter, where they need more time to decipher the rules.

In the L2TOR project, the number as well as duration of lessons we can deliver is limited, and thus giving explicit instructions is expected to be the most efficient way to introduce a new activity. In line with our suggestion, previous research also suggests advantages of explicit instructions over implicit instructions in L2 learning in general (Ellis et al., 2006; Norris & Ortega, 2000). One way to reduce the amount of explicit instructions required is to repeat the same game across sessions. The Turkish teachers used the same set of tasks across different themes. Children became highly familiar with the rules of some tasks, and thus there was no need to re-explain the task in detail. This strategy also gives children an opportunity to contribute to the class by talking about the rules they already know, which may help children be more engaged in the interaction. Therefore, it may be wise to repeat games in L2TOR lessons across different sessions.

2. How do teachers introduce target words for the first time?

In the lessons we observed, almost all words were introduced without explicit translations. Rather, they were introduced via visuals such as flashcards or gestures such as pointing to the relevant object in the classroom. Teacher 2 gave no translation in L1 even in the first session of the lessons. She gave additional information when required (e.g., when children did not understand the difference between two animals). Teacher 1 translated some words when clarification seemed necessary, though very rarely. For instance, when learning names of clothing items (e.g., hat, gloves) using flashcards, one child saw a picture of a glove and said "fingers," and the teacher translated the word

"glove" into Turkish (L1). As the Dutch teachers did not speak the L1 of the children, they could not provide L1 translations of the target words. All teachers made use of cues such as flashcards and gestures to make clear the meaning of each word. Most words taught in these lessons were nouns. As such, guessing the meanings of these words should not be too difficult even for young children (e.g., Gentner, 1981). The L2TOR project may face additional challenges, as it will teach other types or words such as verbs. Supplemental cues such as gestures and direct translations may be necessary when introducing these words, which should be examined through experimental testing. It may be also beneficial to promote metalinguistic awareness (e.g., pointing out similarities between words; "Sitting sounds a bit like walking – they both end with ing!"), partially relying on children's knowledge in L1. The advantage of the robot is that it is able to use the child's L1, allowing for L1 translations across all language settings.

3. What do teachers do when they need children to repeat or clarify themselves?

This question is critical to our project as existing automatic speech recognition (ASR) systems are not yet able to recognize young children's speech with high accuracy. Children's speech is not always clear and is often ungrammatical, and classrooms are often noisy environments. Therefore, we closely looked at occasions in which the human teacher could not understand what the child was saying. In these cases, teachers often tried to reduce the noise in the classroom by warning the students (e.g. "I cannot hear her because it's a lot of noise in here") and then directly addressed the child to speak louder or repeat her answer. It must be noted that teachers did not interrupt children but rather asked for repetition after having finished talking (see Appendix 2 for phrases teachers used in asking children to repeat their words). Within the L2TOR project, children might be shy when talking to the robot, especially in their initial encounters. To initiate and maintain a dialogue, it may be important for the robot to ask children in a natural way to repeat their answers or to talk louder. This would also give children an opportunity to rethink their answer and/or practice their speech.

4. What do teachers do when children say things that are not related to the lesson?

It is not unusual for young children to start talking about a topic that is not directly related to the topic of the conversation that is taking place. Irrelevant talk may be frequent not only in a big classroom where the teacher is not able to pay equal attention to all children at once, but also in a one-on-one lesson if children get bored or do not want to continue the activity. It is observed multiple times in smaller group sessions, and teachers often allow children a limited time to talk before they guide them back to the learning material. This allows children to relieve the urge to express themselves before they refocus themselves on the lesson's contents.

Children will likely talk about irrelevant things in our robot-assisted lesson too, and thus it is fair to ask how professional teachers deal with these irrelevant utterances. The teachers we observed almost always acknowledged such utterances: they listened to the child for a little while and responded to the child. The teachers used topic-specific responses (e.g. "It's not a nose, it's a beak") as well as general confirming phrases (e.g., "Oh, okay"). They also asked the child to wait and talk about the topic later on (e.g., "Later, okay? Don't worry," "Later we will do it"). In either case, after a few exchanges with the child, the teachers quickly went back to their lesson without giving the child another chance to talk about the irrelevant topic. It may be important to implement such

features in order to make children believe that the robot is a conversational partner whom they can really communicate with.

5. What do teachers do to get children engaged?

The lessons we observed were carefully designed to engage young children, for example through playing games. The most prominent engagement activities were singing and doing physical activities such as pretend play, running, or dancing. Teacher 2 used songs for various purposes. She signalled sections of the lesson (e.g., clearing up the toys, concluding the lesson) through songs children were familiar with (e.g., a "Bye-bye song"). This way, not only order was maintained, but also children got the opportunity to practice familiar words in the song. In addition to this purpose, target L2 words were also often introduced and recaptured within a song. These songs were often accompanied with actions and gestures to consolidate the meanings of words (e.g., repeating animal names in L2 while pretending to be the given animal). As children got familiar with songs, they rehearsed the target L2 words by both hearing and pronouncing them. Some of the songs were fully original, while, for some other songs, instructional lyrics were sung to the tune of famous nursery songs. For example, the teacher sang "Make a circle, a make a circle" to the tune of "Ring a Ring o' Roses." Previous research also reports that L2 teachers use chanting and singing with young learners (Fonseca-Mora, 2000). Sounds, rhythms, and intonation patterns in songs can improve children's pronunciation, retention of words, and linguistic structures in an L2 (Foster, 2006). Coyle and Garcia (2014) also found that Spanish-speaking 5- and 6-year-olds learned English words embedded in songs (note, however, that this study did not test a control group that was not exposed to the songs). This project therefore could include songs, although cultural differences must be carefully considered when choosing songs.

As stated above, not only songs but also physical activities engage children into the learning material. The ability to perform actions is one of the unique strengths humanoid robots have. Dancing, pretend-play, and running are a few of the activities we observed children to enjoy in our classrooms. In Classroom 1, children were predominantly asked to sit around a table, which resulted in them getting uneasy throughout the lesson. In Classroom 2, on the other hand, most lesson activities required children to move in the classroom, such as walking around, running, pretending and dancing. Teacher 9 discussed the season of autumn by having the children play with leaves that had fallen off trees (e.g., making the leaves fly, raking leaves). Children in the latter two cases seemed to be more engaged than their peers in Classroom 1.

6. What do teachers do when children are not paying attention?

As described in the earlier section, a variety of verbal (e.g., "Listen!") and non-verbal (e.g., clapping hands) attention grabbers were used by teachers. Teachers, however, used somewhat different strategies to manage their lessons. Some teachers simply gave verbal (e.g., "boys and girls") and nonverbal (e.g., hand gestures, touching the child's arm) warnings and also called individual children's names to have their attention back. Teacher 2, on the other hand, used a song-like warning that alerted children to be quiet in three seconds. The song goes: "Listen, listen to me, one, two, three!". In addition to conventional attention grabbers and warnings, it may be beneficial to develop a song or chant that can be used across sessions, as children are quite good at following a routine. However, it must be noted that the efficiency of attention grabbing strategy (i.e. successfully guiding the child back to the learning material) might differ across

contexts. While chanting might be a better way of drawing attention of numerous children at once, calling individual children's names might be more efficient in quickly drawing a specific child's attention. In the setting of our project where the robot is with a single child, calling the child's name can be the main attention grabbing strategy within a section, whereas chanting can be used to signal the transition to the next section of learning material. Research also suggests that more specific verbal attention grabbers (e.g., "I have a surprise for you") help young L2 learners gain motivation and interest (Mueller & Dweck, 1998). Although non-verbal attention grabbers alone might be difficult or even virtually impossible for the robot to implement (e.g. clapping hands, snapping fingers), their approximations can accompany the verbal attention grabbers to direct children's visual attention to the robot itself when they look away (e.g. waving arm).

7. What do teachers do when children are not answering their question? How long do teachers wait until the child responds?

When the child was not answering the question at all, teachers usually waited for a few seconds. Importantly, the teachers did not simply sit there: while waiting for the child, teachers looked into the child's eyes. If the child seemed to be thinking and about to say something, they waited, but otherwise, they said the answer or gave hints. Similarly, when repeating phrases together with a child, teachers waited until the child looked at the teacher.

If the child was not speaking even after waiting for a while, the teachers often gave hints. The most common hint was the first part of the correct answer (e.g., the first syllable of the word), possibly because most questions demanded the child to say an L2 word or sentence. For instance, Teacher 2 asked the name of an animal on a flashcard (a line drawing of a crocodile in this case), and the child was about to give a wrong answer. Teacher 2 did not even let the child finish the word but elicited the correct answer ("crocodile") by saying the first syllable of the correct answer ("Cro…"; Session 4 Utterance 514). Many children in Teacher 2's classroom remembered names of animals, but they often struggled with forming correct full sentences. For instance, when a child wanted to say, "Rabbit is my favourite animal" but was struggling to finish her sentence, the teacher said "A-, Ani-, Animal", pausing between sounds to provide the child with the opportunity to finish her sentence.

In short, it is important to give children some time to talk while retaining eye contact. When children do not provide an answer, giving hints (both verbally and non-verbally, such as pretending to be the animal if the child tries to remember one) allows them to come up with their own answers.

8. When do teachers use gestures?

Teachers used a variety of gestures for different purposes in their lessons. A list of examples can be seen in Appendix 3. The most common gesture was pointing. Pointing directs children's attention to the relevant object, material or location. It can also be used to further illustrate the meaning of (part of) an L2 utterance, as in the case of pointing to a picture or an object.

In our sessions, the number of gestures varied across teachers as well as across sessions of the same teacher. Such variance might be at least partly due to the lesson theme and teachers' teaching styles. For instance, Teacher 1 used many gestures when teaching body parts (e.g., pointing to her arms), but did not produce many gestures when teaching colours and used flashcards instead. Teacher 2 used fewer gestures

compared to Teacher 1. She was teaching animals across all three sessions and her materials were predominantly flashcards. Based on these observations, it seems that when visual teaching materials such as flashcards are present, teachers mostly depend on them and are less likely to use their hands and body to teach.

One notable pattern we observed was that the teachers used gestures only when there was a conventional or very straightforward gesture associated with the word they were teaching. Though the teachers may have used conventional gestures simply for practical reasons (e.g., it is not easy to come up with original gestures on the spot), this pattern certainly raises the question of whether we should force ourselves to develop novel gestures for words we do not have somewhat conventional gestures for. We must be aware that gesturing too much can be confusing and distracting for young children, as they would have an extra task of deciphering the meaning of the gesture. In addition, in the case of robots, using too many gestures potentially entails different problems such as distracting the child by motor noises. These concerns are currently evaluated in experimental settings.

As discussed above, all gestures carried out by the teachers were very much conventional and straightforward, such as pointing, that even young children are expected to have previously seen. This also held for iconic gestures. However, the use of iconic gestures was very limited, possibly because of the topics covered in these lessons. For instance, Teacher 1 taught colour names and body parts. There cannot be any iconic gestures for colour names, and pointing (which not an iconic gesture) is much easier to do for body parts. Iconic gestures were observed in Teacher 2's lessons, again possibly because her lesson themes allowed their use. When Teacher 2 was teaching names of animals, she used many gestures such as waving her arms when they were singing about the octopus and clawing when she was talking about cats. She also used iconic gestures to remind children the correct answer to her question, such as pinching as a hint for the word crab. The use of iconic gestures was also found in the English lessons observed in the Netherlands. For example, one of the teachers pretended to carry a box as she used the word "heavy" (Session 1). In another session with "the weather' as its main theme, the teacher performed iconic gestures with her hands and whole body when referring to words such as "snow" and "windy" (Session 3). While some iconic gestures might be difficult for our robot to perform, using them whenever the lesson content and robot's structure allow should facilitate the learning process (e.g., Goldin-Meadow, Kim, & Singer, 1999). We are currently conducting an experiment in which the robot is using iconic gestures for animals (e.g., monkey and chicken) to investigate whether those gestures facilitate word learning.

9. When do teachers use the L1?

As noted, the use of the L1 was not possible for the L2 Dutch settings, as the teacher did not speak children's L1. Therefore, we rely solely on the L2 English data for this question. Decision of when and how to use L1 depends on various factors such as teachers' preference and teaching styles and children's language skills. In our observations, Teacher 1 mixed L1 and L2 to get the same message across and to consolidate the meaning of the L2 sentence (e.g. "Can you touch your shoulders? *Omuzlarına dokun*. Touch your shoulders."). When children did not understand an utterance in L2, she gave the verbatim translation in L1. She also used L1 to clarify the meaning of words. For example, in her second session, she showed flashcards with pictures of clothing items (e.g., hat, gloves) and asked children to name them. Children

were generally good at the task. Thus, there was not much need for the teacher to translate the English words. For a picture of gloves, however, one child said "fingers," and the teacher translated the word "glove" into L1. In short, when visual cues were not sufficient or children to deduce the meaning of a word, Teacher 1 translated the word into L1. Apart from utterances that are related to lesson content, she also used L1 to maintain silence in the classroom, to draw children's attention and to give instructions (e.g. "Please count in English").

Teacher 2 adhered to the L2 more strongly. She used L1 only if children asked a personal, mostly health-related question irrelevant to the lesson. Even in that case, she first answered in L2 but switched to L1 only when the child did not understand the answer. In contrast to Teacher 1, moreover, Teacher 2 invited children to use L2 and did not respond to them when they used L1. Thus, she had a strict immersion approach in which she only used L2 and children were only allowed to speak in L2. Whenever children spoke L1, she reminded them to use L2 through phrases like "If you can say it in English, I am ready to listen to you" and "Yes, in English. Oh, that's Turkish. Sorry."

While both approaches might be efficient in L2 teaching, the first approach held by Teacher 1 seems to have short-term benefits (i.e., the child has the immediate chance to associate the L2 verb/noun with the action/object). Research suggests that the use of L1 can provide scaffolding in L2 learning (Mitchell & Miles, 2004). The latter approach, in contrast, might have more long-term benefits in learning (i.e., the child must deduce the meaning of the L2 verb/noun herself. Although it takes more time, this is an example of active learning where the child explores the learning material herself, making her own inferences). Further, other factors may play a role, such as children's home language backgrounds. Specifically, in in lessons teaching immigrant children, using children's L1 is typically not an option as teachers usually do not speak children's L1s.

10. When and how do teachers use positive feedback?

Teachers constantly used a substantial amount of positive feedback (see Appendix 4 for the list of positive feedback used by the two teachers). On average, the teachers provided positive feedback 1.74 times per minute, which translates into 34.83 times for a 20-minute lesson. These feedback comments were often accompanied with nodding. Teacher 2 also used the "thumbs up" gesture frequently. These actions were also observed in L2 English lessons in the Netherlands.

Apart from the list of phrases given in Appendix 4 and gestures such as nodding and giving thumbs up, Teacher 2 acknowledged the correct answer of children by repeating or expanding on their answers. For instance, if the child correctly named the animal on the flashcard as whale, she said "You have a [sic] whale". She also often repeated the correct answer (e.g. Penguin, you have a [sic] penguin"), as further discussed in response to Question 13. The L2 Dutch teachers repeated the words of the child most often as a positive feedback, whenever the child pronounced the target word, the teacher repeated the target word sometimes, accompanied by a praise such as "well done" sometimes only the target word.

11. When do teachers use negative corrective feedback to correct children's answers?

In our sample, the use of negative feedback was considerably less frequent than the use of positive feedback. Still, how teachers handle wrong answers must be carefully addressed. Explicit negative feedback came in the form of negation (e.g. "No", "I don't think so" or "It was not squid") and correction (e.g. "No, that's brown"). It may also be

used in combination with other types of feedback. For instance, on one occasion, the child's response was "[sic] lion" where the correct answer was "seal". The Teacher then first used negation ("No, that's not a sea lion") and then proceeded to explain the characteristics of a sea lion ("A sea lion has a longer tail and the colour is gold"). After this explanation, the child gave the correct answer. Importantly, after that correct response, Teacher 2 further expanded on the description of a seal ("That's a seal, okay? With the short tail and round face"). Providing explanations after negative feedback may help children to learn better. Teachers also used implicit negative feedback, indicating that the answer is wrong without explicitly telling so. For instance, when the child had the flashcard with a jellyfish but could not say what animal it was, Teacher 2 said "You have jellyfish, okay?".

12. Why do teachers repeat some words more than other words?

Some target words were repeated more often than others for a very simple reason, for example, because they occurred in a song with many repetitions (e.g., Teacher 1 sang "Head, Shoulders, Knees and Toes" and thus those four body parts were mentioned more often than other body parts). Other target words were repeated multiple times because children were interested in the topic and the teacher went along with it. For example, the word "cat" was mentioned more often than the word "dog" in Teacher 2's session, and that was simply because one of the children wanted to talk about the name of the cat on the flashcard. Though this organic nature of teacher-child interaction is difficult to replicate in robot-child interactions, it is quite important to build conversations upon what the child says because research suggests that children learn best from what they are interested in (see Konishi et al., 2014 for a review).

The teachers also repeated words that children did not seem to have fully memorized yet. For example, when the child failed to answer a question, the teachers chanted the answer multiple times afterwards. Interestingly, we also observed that Teacher 2 usually repeated the child's answer, even when the answer was correct (e.g., "Zebra is my favourite animal, thank you"). The same strategy was also found in lessons observed in the Netherlands. This appeared to be an effective strategy since, besides being a natural way of responding to children's answers, it enabled all children to hear the correct pronunciation and consolidate their knowledge. Children did not seem to have any problems with hearing the words again. Exposing the child to target words as much as possible is definitely a good strategy, and saying the correct answer no matter what may be an easy and effective way to increase the amount of exposure.

5 Discussion and Conclusion

The semi-naturalistic observation of L2 classrooms provided several useful insights into the teaching strategies as well as phrases and sentences that can be used in our lessons. Our observational data suggest patterns specific to the teacher or lesson topic, as well as common patterns present across teachers and lessons. In the quantitative analyses section of this deliverable, we provided an overview of how our 20-minute robot-assisted sessions should look like if we were to mimic the way human adults teach L2 (as based on the current observations). Notable characteristics we found in the quantitative analysis were the minimal use of L1 and negative feedback, which must be carefully considered in developing our own lessons. In the descriptive analysis section, we attempted to answer practical and theoretical questions that may arise as our

robot and lesson plans are developed further. To note a few, we found that the use of gestures was highly context dependant and singing songs and physical activities (e.g., dancing) were used frequently across lessons to engage young learners.

It must be noted, however, that there may be a limit on how much of what we observed can be and should be adapted in robot-assisted lessons. For example, although we found that the teachers constantly performed actions and gestures to motivate children and to facilitate their learning process, the robot gesturing too much might cause more harm than good, especially because gestures of our robot will not be as smooth as that of humans and because the motor sounds the robot produces while gesturing may be quite distracting. Semi-naturalistic observations can give us an idea, or the starting point, about what may be suitable for our lessons, but cannot inform us about the effects of particular teaching strategies. For example, although we observed various kinds of feedback in the L2 lessons, on the basis of our data, we cannot draw conclusions on their effectivity in robot-assisted L2 tutoring. Further, it is important to note that the observations were based on L2 classroom sessions, whereas the lessons to be developed in the L2TOR project involve individual (child-robot) lessons.

The data presented in this deliverable provide a first starting point for developing lessons within robot-assisted L2 learning, which needs further refinement and validation through experimental evaluation. We also expect more questions to arise as our robot is designed further because it is impossible to foresee all possible issues at this point. Thus, we plan to conduct additional coding of the observation data in concert with the development of the robot. Experimental studies are currently being undertaken within several WPs of the L2TOR project and will be presented in future deliverables.

References

Azaoui, B. (2013). One Teacher, Two Instructional Contexts. Same Teaching Gestures? In *TiGeR: Tilburg Gesture research meeting,* 1-4.

Coyle, Y., & Gracia, G. R. (2014). Using songs to enhance L2 vocabulary acquisition in preschool children. *ELT Journal*, *68*(3), 276-285.

de Haan, A. K. E. (1982). Effects of preschool education in mixed and targeted classrooms (Doctoral thesis, Utrecht University, Utrecht, the Netherlands). Retrieved from https://dspace.library.uu.nl/handle/1874/322447

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28*(2), 339–368.

Fonseca-Mora M. C. (2000). Foreign language acquisition and melody singing. *ELT Journal*, *54*(2), 146–152.

Forster, E. (2006). The value of songs and chants for young learners. *Encuentro, 16*, 63–8.

Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory, 4*(2), 161-178.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teachers' hands tell the students' minds about math. *Journal of Educational Psychology, 91*, 720–30.

Goldin-Meadow, S., & Sandhofer, C. M. (1999). Gesture conveys substantive information to ordinary listeners. *Developmental Science, 2,* 67–74.

Gullberg, M. (2008). Gestures and second language acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 276-305). London, UK: Routledge.

Kanda, T., Shimada, M., & Koizumi, S. (2012). Children learning with a social robot. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 351–358). ACM.

Konishi, H., Kanero, J., Freeman, M. R., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Six principles of language development: implications for second language learners. *Developmental Neuropsychology, 39*, 404–420.

McCafferty, S. G. (2002). Gesture and creating zones of proximal development for second language learning. *Modern Language Journal*, *86*, 192–203.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

Mitchell, R., & Myles, F. (2004). *Second language learning theories (2nd ed.)*. London, UK: Hodder Arnold.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*(1), 33–52.

Muramoto, N. (1999). Gesture in Japanese language instruction: The case of error correction. In L. K. Heilenmann (Ed.), *Research issues and language program direction* (pp. 143-175). Boston, MA: Heinle & Heinle.

Noor, N. M., Aman, I., & Mustaffa, R. (2012). Teachers' Questioning Approaches in the Malaysian ESL Classroom. *International Journal of Learning*, *18*(7), 313-326.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*(3), 417–528.

Taleghani-Nikazm, C. (2008). Gestures in foreign language classrooms: An empirical analysis of their organization and function. In *Selected proceedings of the 2007*

second language research forum (pp. 229-238). Somerville, MA: Cascadilla Proceedings Project.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, *8*(2), 219-235.

Yao, S. (2000). *Focus on form in the foreign language class- room: EFL college learners' attitudes toward error correction*. Buffalo: State University of New York at Buffalo.

Appendix 1: Coding Scheme

We coded **5 main components** of a typical L2 teaching/learning context.

1. **Action/gesture use**
   a. Target word related
   b. Non-target word related
2. **Attention grabbers**
   a. Nonverbal
   b. Verbal
   c. Effectiveness
3. **Interaction patterns**
4. **Feedback patterns**
   a. Affective feedback
   b. Corrective feedback
      i. Positive feedback
      ii. Negative feedback
5. **Language use**
   a. L1
   b. L2
   c. Intra-sentential code-switch
   d. Inter-sentential code-switch

---

1. **Action/gesture use**
   We code teacher action/gesture use in two categories:
   The videos we have mostly lend themselves for action or deictic gesture use code. Of course, gesture use depends on individual preferences and the teaching context/theme. We have coded for specific units for action/gesture use for the time being. We will eventually form main categories out of these specific single units.

   a. Target word related action/gesture use:
   Target words are those used in that specific session.
   e.g. If the session is about clothes, then you would expect target words to be clothing items such as scarf, hat, gloves, shoes and jacket. Examples are below:

   - showing
     e.g., showing a flashcard (which has a drawing of a jacket-target word)
   - looking
     e.g., looking at a flashcard (which has a drawing of a jacket-target word)
   - moving
     e.g., fingers to gesture a target word- gloves
   - pointing
       i. index finger pointing (e.g., pointing at an object *or* at a specific location)
      ii. whole hand pointing (e.g., pointing at an object such as a jacket)
     iii. body pointing (e.g., touching head while teaching the word 'head')
   - pretend-play
     e.g., pretending wearing a jacket

b. Non-target word related action/gesture use:
   Non-target words include any utterance (L1, L2 or a mixture of L1 and L2) used in class including giving feedback, explanations or engaging in any sort of interaction. Examples are below:

- hand gesturing
  e.g., gesturing with fingers to show number 2 while instructing children to work on two flashcards
  e.g., gesturing with both hands to instruct children to go back
- pointing to a location

2. Attention grabbers
   These are single words or phrases or non-verbal gestures used by the teacher to draw the attention of the children when they lose motivation or concentration during an activity. Examples are below:

a. Nonverbal examples:
- snapping fingers while asking an individual child to respond
- pointing at a specific child
- clapping
- touching (the head of) a child
- looking at a specific child
- knocking on the table
- pausing

b. Verbal examples:
- (saying) 'listen!'
- (saying) 'look!'
- calling out on an individual child

c. The effectiveness of attention grabbers:
   This is rated on a scale: 0 (ineffective), 1 (somewhat effective), 2 (very effective)

3. Interaction patterns
   As an instructional methodology, these are any verbal strategy the teacher uses to conduct the class activities. Some examples are below:

- making analogy (e.g., the sound 'le' in legs is resembled to the pronunciation of another related word in L1)
- self-correction (e.g., yellow hair# oh sorry# blonde hair)
- responding to an L1 question in L2
- giving instruction (e.g., everybody please colour gloves, can you turn two cards?)
- asking a question (e.g., what colour is it?)
- repeating instruction/question
- asking for individual or choral repetition
  (e.g., Child's name# can you repeat it, everybody repeat it)
- repeating with child(ren)
- asking for clarification (e.g., why?)

- asking for silence (e.g., be quiet! ssshhh!)
- giving an explanation
- verbatim translation (e.g., word by word translation from L1 to L2 or vice versa)
- exaggerated pronunciation of certain words or sounds
- chanting (e.g., singing a theme related song);
- directing to L2 (e.g., in English please)
- commenting and evaluating (on an activity)
- describing a picture (e.g., yellow gloves)
- giving the objective of the class (e.g., Today we will learn clothes)
- signalling a new task (e.g., Let's start, are you ready?, now we will play one game)
- giving a clue (e.g., what do we put on our hands?)
- concluding a task (e.g., that's it)
- giving permission (e.g., you can take your crayons)
- exemplifying (e.g., like that)
- saying farewell (e.g., ok good bye everyone, see you tomorrow)
- introducing a new word (e.g., gloves) or a theme (e.g., winter clothes)
- calling out on an individual child (to take turns in doing the activity)
- thanking (e.g., thank you)
- warning (e.g., don't sleep, sit properly)

4. Feedback patterns
These include affective and corrective feedback given by the teacher:

a. Positive feedback:

  i. Explicit examples:
- Praise (e.g., good job, well done, good boy, good boy)
- Encouragement (e.g., you can do it)
- Confirming (e.g., yes, that's it!)

  ii. Implicit examples:
- Consolidation (repeating the accurate response of the child)

b. Negative feedback:

  i. Explicit examples:
- Correction: an overt and clear indication of the existence of an error and the provision of the target-like reformulation
(e.g., No, this is wrong, it is not a cat, it is a dog)
- Negation: Just a "no," or "no, that's wrong."

  ii. Implicit examples:
- Recast: teacher's reformulation of all or part of a student's utterance, minus the error.
- Expansion: correcting the child and adding on what the child says
- Clarification requests: questions indicating that the utterance has been ill-formed or misunderstood and that a reformulation or a repetition is required.
- Repetition: teachers' repetition of the ill-formed part of the student's utterance, usually with a change in intonation.

5. Language use

This refers to the language (L1, L2 or code switching) used by the teacher. Coding for language use includes language used while doing a class activity and/or dealing with a problem in class.

a. L1 use: Teacher's use of first language (L1) in the teaching context
b. L2 use : Teacher's use of second language (L2) in the teaching context
c. Intra-sentential code-switch use: Teacher's interjection of a phrase, word or a single sentence from one language into the other (e.g., "*Ich habe shoes*," "*hadi simdi gloves'i boyuyalim*"). This relates to the use of two different languages <u>within the same utterance</u>. Words belonging to different parts of speech such as nouns or verbs are uttered in L2 and they are inflected with L1 morphology).
d. Inter-sentential code-switch use: This relates to the use of two different languages across utterances (e.g., "*Ja*, I like it," "shoe, shoes, *iki tane olunca çoğul kullaniyoruz*").

Appendix 2: List of phrases teachers used to ask the child to repeat

- Let's repeat it [child's name]
- I cannot (really) hear you
- Sorry, I couldn't / cannot hear (you)
- Repeat it
- Please loud
- Can you repeat it?
- One more time

Appendix 3: List of purposes and examples
of teachers' gesture use

- To clarify a target word/phrase:
  Example 1: pretending wearing shoes while saying "put on your shoes." (Teacher 1)
  Example 2: pointing to herself while saying "my favourite." (Teacher 2)

- To clarify the instructions :
  Example 1: showing two fingers while saying "turn two cards." (Teacher 1)
  Example 2: shaking her hand back and forward while saying "go back." (Teacher 1)
  Example 3: shaking her hand up and down while saying "stand up." (Teacher 2)

- To draw children's attention to an object that is related to current activity:
  Example 1: pointing to a laptop while saying "now we'll watch a video." (Teacher 1)
  Example 2: pointing to a board while saying "I'll write your points to [sic] the board." (Teacher 2)

- To add information to clarify the instruction:
  Example 1: pointing to the floor while saying "put the flashcard here." (Teacher 2)
  Example 2: stop gesture with hand while saying "wait for the question." (Teacher 2)

Appendix 4: List of positive feedback phrases

- Çak (high five)*[1]
- Alright
- Bravo
- Clever boy
- Come on
- Good
- Good boy/girl
- Great
- Hah
- Ha-ha
- Hı-hı*[2]
- Hııh*[3]
- It's alright
- Okay
- Thank you
- That's it
- That's alright
- That's great
- Very good
- Very nice
- Well done
- Yes/yeah
- You are (a) very clever girl
- You got it

*1: "Çak" means "high five" in Turkish
*2: "Hı-hı" is an expression used to show assent in Turkish. It is similar to the expression "uh huh."
*3: "Hııh" is an interjection in Turkish to show approval, similar to "aha" or "there."