CoWriting Kazakh: Transitioning to a New Latin Script using Social Robots

Anton Kim[1], Meruyert Omarova[1], Adil Zhaksylyk[1],
Thibault Asselborn[2], Wafa Johal[2,3], Pierre Dillenbourg[2], and Anara Sandygulova[1,*]

***Abstract*— In the Republic of Kazakhstan, the transition from Cyrillic towards Latin alphabet raises challenges to teach the whole population in writing the new script. This paper presents a CoWriting Kazakh system that aims to implement an autonomous behavior of a social robot that would assist children in learning a new script. Considering the fact that the current generation of primary school children have to be fluent in both Kazakh scripts, this exploratory study aims to investigate which learning approach provides better effect. Participants were asked to teach a humanoid robot NAO how to write Kazakh words using one of the scripts, Latin vs Cyrillic. We hypothesize that it is more effective when a child mentally converts the word to Latin in comparison to having the robot perform conversion itself. The findings reject this hypothesis, but further research is needed as it is suggested that the way the pre-test was performed might have caused the obtained results.**

***Index Terms*— Human-Robot Interaction, Robots for Learning, Language learning, Social Robot, NAO**

I. INTRODUCTION

Kazakhstan has recently adopted a state program for the development and functioning of languages for 2011-2020. This new trilingual education policy aimed at development among the Kazakhs of fluency in three languages: Kazakh, Russian and English. Additionally, a recent decision on the transfer of Kazakh language from Cyrillic into the Latin alphabet was approved by the Kazakh authorities in October 2017 [1]. While there are clear reasons for these reforms, there are numerous risks facing the transfer, including risks to cause disinterest and lack of motivation to learn the Latin-based Kazakh among school children.

A substantial increase in social robots in various areas of applications raises the importance of human-robot interaction research, especially in the application of education [6]. Recent years have seen the increase in the amount of research investigating the topic of language learning using social robots [7], [10]. Numerous advantages make social robots effective for such applications, for example a possibility of adapting to child's language level during a vocabulary learning task [8].

Since 2014, the CoWriter project has explored how robotic technologies can help children with the training of handwriting via an original paradigm known as learning by teaching (LbT) [2] [3] [4]. Since the children act as the teachers who help the robot to learn handwriting, the children practice their handwriting even without noticing it and stay committed to the success of the robot via the Protége effect. Previous research have shown the motivational aspect of the LbT with a robot for handwriting [3]. We believe that the CoWriter activity has the required innovative aspect to it and, hence, it can boost the childrens self-esteem and motivation to learn the Latin-based Kazakh alphabet and its handwriting. This paper presents the CoWriting Kazakh project that aims to benefit from the new language planning in Kazakhstan in order to address challenges of training and motivating children to learn and use a new alphabet.

In the CoWriting Kazakh system, the NAO robot is introduced to a child as a native English speaker that needs child's help in learning new vocabulary in Kazakh language. The robot asks the child to translate simple words from English to Kazakh (e.g. “hello”) and to demonstrate how to write them using Latin script so that the robot can read it. To investigate whether this approach would cause children learn more letters in Latin-based Kazakh, we conducted an exploratory study with 48 children in primary school in Kazakhstan. Children were asked to complete a pre-test, robot interaction, and a post-test to evaluate the number of learned letters in Latin-based Kazakh. Children interacted with the robot in one of the conditions where they had to demonstrate how to write Kazakh words using either a Cyrillic script or a Latin script. The robot then had to repeat child's demonstration using only Latin script that was a correct spelling of the words so that the child could see the right spelling in Latin script and learn from his/her mistakes. We hypothesize that Latin demonstration by the child (in contrast to Cyrillic demonstration) is more effective for learning a new script.

II. HRI SYSTEM

Since the project is motivated by the recent decision of Kazakh authorities to transition from Cyrillic to Latin script [1], we aim to train students of a new script in a learning scenario with a social robot. In contrast to original CoWriter's LbT paradigm where robot's handwriting improved gradually via several demonstrations by the child, the CoWriting Kazakh does not have a handwriting improvement component. In the presented system, the robot and a child engage in a co-operative learning where the robot learns from the child the new vocabulary in Kazakh while the child learns

*Corresponding author: anara.sandygulova@nu.edu.kz
[1]Department of Robotics and Mechatronics, Nazarbayev University, Kabanbay Batyr Ave 53, Nur-Sultan, Kazakhstan
[2]CHILI Lab, École Polytechnique Fédérale de Lausanne, Route Cantonale, 1015, Lausanne, Switzerland
[3]BIOROB Lab, École Polytechnique Fédérale de Lausanne, Route Cantonale, 1015 Lausanne, Switzerland

Fig. 1. Screenshot of a Wacom tablet demonstrating Cyrillic-to-Latin condition

from the robot the spelling in new script. Thus, they turn turns in writing words in Kazakh (see Figure 1).

A. *Robot Role*

In this scenario the NAO robot plays a role of a peer. The robot is introduced to a child as a native English speaker of approximately his or her age who wants to learn Kazakh. The robot asks the children for help, especially, to demonstrate how to write Kazakh words using the new Latin alphabet because that is convenient for the robot to read. In a control condition, the robot does not ask to write explicitly in Latin script, so the child writes words in their preferred script which is Cyrillic script as that is what they are used to.

The child is told that the robot does not understand Kazakh or Russian language, so children have to understand what the robot says. The reason is the absence of the child voices in Kazakh text-to-speech engine. It was important to compose simple robot speech utterances for the children to understand. We developed these utterances and then verified them with the help of the English teacher.

B. *Dialogue*

Interaction with the child consists of several stages during which the robot sustains the interaction: greets the child in the beginning, provides instructions in the form of the questions, and says goodbye at the end.

NAO: -Hello. I am a robot. My name is Mimi. [Waves his hand]
Child: -...
NAO: -I study Kazakh language. Can you help me?
Child: -...
NAO: -How do you say “Hello” in Kazakh?
Child: -Sálem
NAO: - How do you write it? [In Latin-to-Latin case: Please write it using Latin letters so that I can read it.]
Child: -[Writes on a tablet the word in one of the scripts]
NAO: -Let me try to write it too [gesticulates]. This is a correct writing using Latin letters.

NAO: -How do you say “thank you” in Kazakh?
Child: -Rahmet
NAO: -How do you write it? [In Latin-to-Latin case: Please write it using Latin letters so that I can read it.]
Child: -[Writes on a tablet the word in one of the scripts]
NAO: -Let me try to write it too [gesticulates] This is a correct writing using Latin letters.

... repeated for another 9 words for a total of 11 words

NAO: - You are a great teacher. Thank you very much! Goodbye! [waves]

C. *Software and Hardware components*

The Wacom Cintiq Pro tablet is a graphics tablet which can display the monitor of the computer, serving as the second monitor. Its pen has 8,192 levels of pressure sensitivity and tilt recognition. This allows to acquire not only the trajectory of the handwriting, but also the pressure and tilt at every point.

A humanoid robot NAO is a programmable autonomous robot developed by SoftBank Robotics. It is widely used in human-robot interaction research, in particular, educational and robot-assisted therapy applications. A humanoid robot’s height is 58 cm which makes it comfortable to transport, also its appearance is appealing for children. Furthermore, it has 25 degrees of freedom and 7 tactile sensors. We extended the original CoWriter project[1].

If a child believes that there is mistake in his/her writing they can use “Eraser” icon, which will clear the writing space. When a word is finished, a child needs to click“Done” icon and a robot will write the same word in the Latin script. After the robot is done writing it will ask for translation of a next word. If children felt tired or bored, they could finish interaction whenever they wished by pressing “Finish” icon (black-and-white flag). There were eleven English words asked by the robot in total.

D. *Conditions*

In order to investigate whether it is more effective for the child to perform conversion mentally and observe correctly written Latin spelling by the robot, we distinguish two conditions that are different in who performs the conversion:

- Latin-to-Latin: the child does the conversion mentally and writes directly in Latin.
- Cyrillic-to-Latin: the robot does the conversion. The child writes in Cyrillic and observes the Latin writing provided by the robot.

During the interaction we did not help children in writing and did not correct their mistakes. We would only help them in case they did not understand or hear the robot.

[1] https://github.com/chili-epfl/cowriter

III. HRI Study

A. Recruitment

This research was approved by the Ethical committee of Nazarbayev University. We commenced an introductory sessions with two classes of different age groups where we gave a brief description of the research. Assent and consent forms were distributed and children were able to ask questions. Teachers then collected assent and consent forms for us in the next few days.

B. Procedure

The procedure of the experiment had the following parts: pre-test, pre-interaction survey (further pre-survey), learning activity, post-interaction survey (further post-survey) and post-test (see Fig. 3). Overall procedure for one child lasted approximately 20 minutes.

Each child was called out of the class and walked with the first researcher for approximately two-three minutes to the room. While walking with the child, the first researcher started with an icebreaker warm-up talk necessary to relax and engage the youngster. "My name is Meruyert and what is your name?", "Have you even seen a robot before?", "When I was in school, I liked Mathematics and what is your favourite subject? After entering the room with the robot, children were invited to take a seat at the table with questionnaires and answer a few questions about their age, gender, and mood prior to the interaction with the robot. Then, children were asked to seat next to Experimenter 2 to complete a pre-test to find out if children know Latin-based Kazakh alphabet. After the questionnaire and tests were filled in, children were invited to change the table and take a seat facing the robot. After the interaction, children were interviewed by the first researcher who followed a structured interview about the interaction. During the post-test the child had to attempt the letters that were incorrectly written during the pre-test. In the end, the first researcher brought the child back to the class and called out the next participant.

C. Pre-survey

Upon arrival to the room, a child would pass a pre-survey (transition 1 on Fig. 3) - a small question-answer session, during which experimenter would record child's name, surname and mood. The following is a list of the questions:

1) What is your name?
2) How old are you?
3) What class are you in?
4) What is your grade on Kazakh language subject? (range 2-5)
5) What is your academic performance? (range 2-5)
6) What is your mood? (a 5-Likert scale)

Experimenter would paraphrase questions to use unofficial language to make sure children feel at ease.

D. Pre-test

The next phase was a pre-test (transition 2 on Fig. 3), where a child was asked to write letters of Kazakh alphabet firstly in Cyrillic script and then in Latin script. The pre-test was needed in order to determine child's level of knowledge of Latin script. It was decided to choose difficult letters to convert or remember for the test, and thus, the test included twenty three lower case letters from 42 and 32 Cyrillic and Latin alphabets respectively. A child was presented with a letter in Cyrillic script and asked to write it, as shown in Fig. 2(a). Pressing "Next" icon (green arrow at top-right corner) opened a next screen (Fig. 2(b), where the child was asked to write the same letter in Latin script. If the child wrote the letter correctly, the second experimenter continued with the next Cyrillic letter by pressing the "Next" icon. But if the child wrote the letter incorrectly or if he/she did not know how to write the letter, the second experimenter pressed the "Show" icon, which displayed the correct writing of the letter in Latin script, and asked the child to write it (see Fig. 2(c)).

E. Activity with the robot

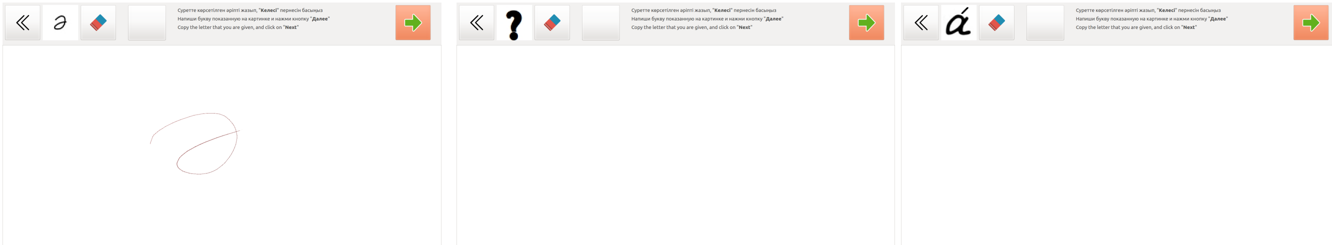
The child was asked to take a seat in front of the robot (transition 3 on Fig. 3). The scenario and a dialogue are discussed in Section 3.

F. Post-survey

After the main interaction there was a post-survey (transition 4 on Fig. 3) where a child was asked a set of questions.

1) What is your mood? (a 5-Likert scale)
2) Funometer [11]
3) What is the robot to you? (a multiple choice question: a toy, an electrical appliance, a computer, a human, or a pet)
4) Who is the robot to you? (a multiple choice question: a friend, a parent, a sibling, a classmate, a teacher, or a stranger)
5) What is the most/least interesting? (Sorting is recorded on a 5-Likert scale)
6) What is the most/least effective? (Sorting is recorded on a 5-Likert scale)
7) Which one do you like the most/least? (Sorting is recorded on a 5-Likert scale)
8) What is the easiest way to learn with/from? (Sorting is recorded on a 5-Likert scale)

The questions' aim was to determine the extent to which children liked the interaction, moreover, we asked what was their mood after the procedure and how they perceived a robot. In more detail, these questions were oriented on examining the inner states of children and various techniques were used to deliver questions to be understandable as possible. For example, the Funometer scale and a picture of a NAO were printed on a paper for children to physically drag the printed robot and position it on a scale. It proved to be more appropriate than having standard 5-Likert scales, since majority of children placed the robot near 80-90 percent instead of 100.



(a) Child is shown a letter in Cyrillic script (b) Child is asked to write the letter's analog in Latin script (c) Child made mistake or does not know a correct answer, so the help icon was pressed

Fig. 2. Pre-test GUI.

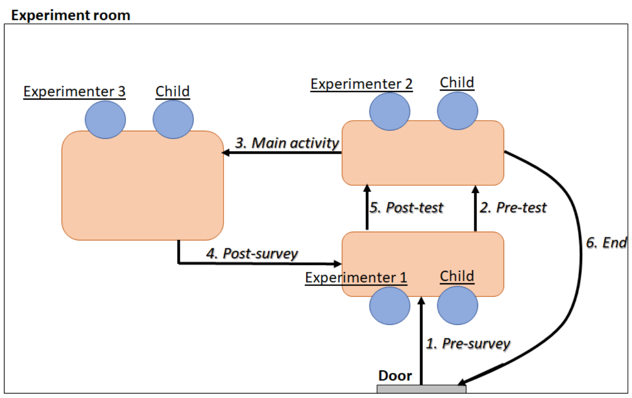


Fig. 3. Schematic flowchart of the experimental procedure.

G. Post-test

The last phase of the procedure was the post-test (transition 5 on Fig. 3). During the post-test children were asked to write Kazakh letters in Latin script again, but only those letters, which they didn't know or made mistakes during the pre-test. The interface was similar to the pre-test, with the difference that there was a screen with Cyrillic version of a letter as in Fig. 2(a) and they needed to write Latin version of the letter on the same screen. Pressing "Next" icon would move to the next letter. However, if a child did not remember the correct writing of the letter or if experimenter noticed that the child made a mistake, the "Show" icon was pressed, then a screen similar to Fig 2(c) appeared and a child rewrote the letter in Latin script. The post-test was needed in order to determine the number of learned letters after the interaction with the robot. After the test, children received a book for their participation.

During the experiment were used three laptops, two WACOM tablets, one Apple tablet, a Wi-Fi router, and a NAO robot (Fig. 5). One set of laptop-WACOM was used by the Experimenter 2 for the pre- and post-tests, and another set of laptop-WACOM was used by the Experimenter 3 for the main activity. The third laptop and the Apple tablet were used by Experimenter 1 for the pre- and post- survey. This way we had a conveyor approach (see Fig. 5), interacting with up to three children simultaneously and, thus, saving a lot of time.

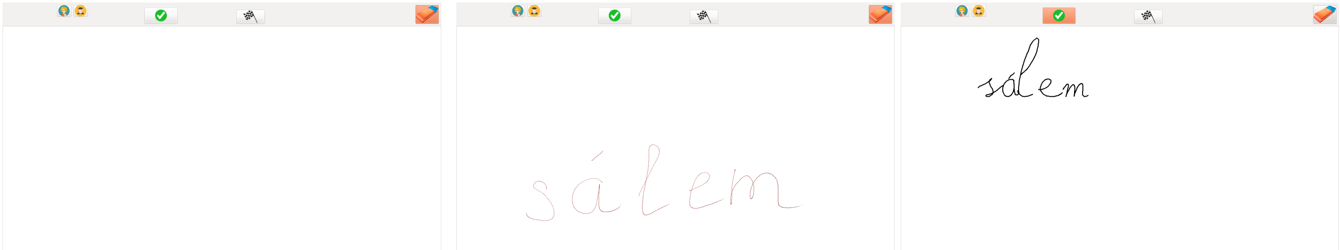
H. Measurements

Data was collected from the questions and performance during the interviews prior, during and after the interaction with the robot.

- **Personal information.** Each child was telling his/her personal information: name, gender, age, class grade.
- **Academic performance.** Children were asked about their grade in Kazakh language and overall academic performance on a 5-Likert scale.
- **Change in Mood: pre- and post-mood.** Children were asked to rate their mood before and after the interaction with the NAO robot.
- **Funometer.** Child was asked to rate how much they liked interacting with the robot from 0 to 100 on a Funometer scale.
- **Robot Type.** The question was about child's comparison of the NAO robot via a forced-choice question: *Toy, Electrical Appliance, Computer, Human, Pet*
- **Robot Role.** Children were asked to compare whether the robot is similar to one of a forced-choice options: *Friend, Parent, Brother/Sister, Classmate, Teacher, Stranger*.
- **Sorting Robot.** Children were asked to physically sort five small pictures (NAO robot, Book, Phone, Teacher, Computer) in a ascending order according to their 1) Effectiveness to teach, 2) Easy to learn from, 3) Interesting, and 4) Enjoyable. We then noted the order (as a 5-point scale) they placed the robot in.
- **Time per word**. While working with the robot, each child's time was recorded. We also recorded how many words each child chose to complete. Thus, we were able to calculate time that it took writing one word on average.
- **Number of words.** There were a total of eleven words asked by the robot to translate and write. However, each child was free to stop at any word.
- **Number of learned letters.** Number of learned letters was calculated after the post-test.

IV. Results

There were 48 children (24 females) aged 8-10 years old. Exactly half of the participants had Latin-to-Latin case which were selected randomly and counterbalanced for gender and age groups. Two children had to leave early and did not



(a) Window right after opening (b) Example of child's writing (c) Example of robot's writing

Fig. 4. Main interaction window.

Fig. 5. Experiment with three children simultaneously

answer post-survey questions. For this reason some of the following tests were conducted on data from 46 children.

A. Task Performances

In general, children improved their knowledge of Latin alphabet during the experiment. The average number of new learned letters is 3.73 (SD = 2.15, Max = 9, Min = 0). Strong learning gains were not observed as we asked for a subset of all letters (23 out of 42 letters). This could be improved by including all letters in the pre-test. This can be verified by performing further research.

A series of one-way ANOVA tests was conducted to determine if there was any difference in learning. It revealed a non-significant difference in the number of learned letters between different robot conditions, which rejects our hypothesis that Latin-to-Latin condition is more effective in co-operative learning scenario.

Then, we conducted a series of one-way ANOVA tests to investigate other dependent metrics. There was a statistically significant difference in performance time between groups as determined by one-way ANOVA ($F(1,46) = 21.94$, $p < 0.001$). A Tukey post-hoc test revealed that the time to complete the task was statistically significantly lower in Cyrillic-to-Latin (11.88 ± 2.54 min) compared to the Latin-to-Latin (15.96 ± 3.43 min). Additionally, there was a statistically significant difference ($F(1, 47) = 13.93$, $p = 0.001$) in children's speed (time per word) between robot conditions: Cyrillic-to-Latin (1.21 ± 0.41 min) compared to Latin-to-Latin (1.83 ± 0.69 min). These significant findings of time differences is expected as it was easier for children to write using a script which children use every day.

B. Gender differences

There was a statistically significant difference in the number of learned letters between boys and girls ($F(1, 46) = 4.03$, $p = 0.05$). Boys learned significantly more (4.33 ± 2.24) letters than girls did (3.13 ± 1.92).

In addition, while an average time per word for all children was 1.52 minutes, there was a statistically significant difference between boys and girls: ($F(1, 46) = 8.76$, $p = 0.005$). Boys spent more time per word than girls did: 1.77 minutes per word vs 1.27 minutes per word respectively.

Looking at time and number of words separately, there is an insignificant difference between time spent for males and females: $t(46) = 1.790$, $p = 0.080$ and significant difference in number of words: $t(46) = -2.467$, $p = 0.017$. Boys spent 14.83 ± 3.83 minutes and did 9.17 ± 2.28 words while girls spent 13.00 ± 3.24 minutes and completed 10.46 ± 1.18 words on average. These findings suggest important gender differences: boys' speed was slower and they chose to complete less words than girls did, but they learned more words in comparison to girls.

C. Academic level differences

Children answered with three categories to the questions on their academic performance: A (N=23), B (N=20) or C (N=3) students. A further exploratory analysis of one-way ANOVA identified speed (time per word) differences between children's performance (satisfactory, good, excellent) in school: $F(2, 43) = 6.029$, $p = 0.005$. Results indicate that A students spent the least time per word, 1.23 minutes per word, while B students spent 1.83 minutes. There were only three students with a C academic level, but they spent 1.24 minutes per word.

D. Age differences

We had 3rd and 4th graders in the experiment, thus we grouped their age to two classes according to their grade. There were no significant differences between age groups, apart from time per word was significantly ($p < 0.05$) longer for older children in comparison to younger children: 1.70 vs 1.32 minutes per word respectively. Interestingly, younger children rated the robot significantly more convenient/comfortable than older children did, 3.87 vs 3.08 (out of 5): $F(1, 46) = 5.87$, $p < 0.05$.

E. Robot perception

A series of chi-square tests of independence was conducted to examine the effect of categorical variables (gender or robot condition) on children's Mood Change, Robot Type, Robot Role. We did not find any statistical significant results between boys and girls for these measurements. However, there was an interesting result (though insignificant) between robot conditions in how children responded to Robot Type: x^2(3, N = 47) = 6.759, p = .0.08. In Latin-to-Latin condition, noone compared the robot to a Toy, while there were 5 children who made this comparison in Cyrillic-to-Latin condition. There were 15 vs 9 children who compared the robot to a computer in two conditions respectively.

V. Discussion and Limitations

All participants were from the same school and we can not generalize to confidently say that the same result will be valid in other schools in Kazakhstan.

Another limitation is in the way the pre-test was performed. When children made mistakes in the pre-test, they were shown the correct answers by the experimenter. It was a way of learning these letters from the pre-test, thus we could argue that it probably caused the insignificant differences between the two developed conditions for human-robot interaction.

It is worth mentioning that children's performance was significantly different in two robot conditions. And since there were significant differences between gender groups: boys spent more time than girls and also learned more letters, we need to perform further research with a similar experimental design but with a different pre-test. It should include all 42 letters of Cyrillic Kazakh alphabet without showing the correct answers in cases of mistakes or hesitations.

VI. Conclusion

This paper describes an exploratory study that aimed to investigate whether children need to try to perform conversion of the letters themselves and then see the results in comparison to having conversion performed by the robot. Future work will include creating an automatic optical character recognition and spellchecking embedded in the application so that the child can choose which words to teach the robot himself/herself. Also, it could be worth integrating original CoWriter's handwriting learning to be able to have several demonstrations of the same word to the robot.

ACKNOWLEDGMENT

We wish to thank the primary schools and, in particular, all the children and their parents who participated in this study. This research was sponsored by the UNIGE Leading House for the Bilateral Science and Technology Cooperation Program for the CIS Region under grant number SFG 1038.

References

[1] E. Altynsarina: Kazakhstan adopts new version of Latin-based Kazakh alphabet. 2018. NATION. [Accessed on 9-July-2018: https://astanatimes.com/2018/02/kazakhstan-adopts-new-version-of-latin-based-kazakh-alphabet/]

[2] D. Hood; S. Lemaignan; P. Dillenbourg: The CoWriter Project: Teaching a Robot how to Write. 2015. 2015 Human-Robot Interaction Conference, Portand, USA, March 3-5, 2015. DOI : 10.1145/2701973.2702091

[3] A. Jacq; S. Lemaignan; F. Garcia; P. Dillenbourg; A. Paiva : Building Successful Long Child-Robot Interactions in a Learning Context. 2016. 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, NEW ZEALAND, MAR 07-10, 2016. p. 239-246.

[4] S. Lemaignan; A. D. Jacq; D. Hood; F. Garcia; A. Paiva et al. : Learning by Teaching a Robot: The Case of Handwriting; Robotics and Automation Magazine. 2016. DOI : 10.1109/Mra.2016.2546700.

[5] Deanna Hood, Severin Lemaignan, and Pierre Dillenbourg. When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 15, pages 8390, New York, NY, USA, 2015. ACM.

[6] Mubin, Omar, Catherine J. Stevens, Suleman Shahid, Abdullah Al Mahmud, and Jian-Jie Dong. 2013. "A Review of the Applicability of Robots in Education". Technology For Education And Learning. doi:10.2316/journal.209.2013.1.209-0015.

[7] Balkibekov, K., Meiirbekov, S., Tazhigaliyeva, N., and Sandygulova, A. (2016, August). Should robots win or lose? Robot's losing playing strategy positively affects child learning. In Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on (pp. 706-711). IEEE.

[8] Kory, J. J. M. (2014). Storytelling with robots: Effects of robot language level on children's language learning (Doctoral dissertation, Massachusetts Institute of Technology).

[9] Asselborn, Thibault, Thomas Gargot, ukasz Kidziski, Wafa Johal, David Cohen, Caroline Jolly, and Pierre Dillenbourg. 2018. "Automated Human-Level Diagnosis Of Dysgraphia Using A Consumer Tablet". Npj Digital Medicine 1 (1). doi:10.1038/s41746-018-0049-x.

[10] Belpaeme, T., Vogt, P., van den Berghe, R., Bergmann, K., Gksun, T., de Haas, M., ... and Papadopoulos, F. (2017). Guidelines for designing social robots as second language tutors. International Journal of Social Robotics.

[11] P. Markopoulos, J.C. Read, S. MacFarlane, and J. Hoysniemi. Evaluating Childrens Interactive Products: Principles and Practices for Interaction Designers. Interactive Technologies. Elsevier Science, 2008.