Nordic Journal of Digital Literacy

Universitetsforlaget

VITENSKAPELIG PUBLIKASJON

Volume 18, No. 2-2023, p. 100–115
ISSN online: 1891-943X
DOI: https://doi.org/10.18261/njdl.18.2.3

Playing with Pronunciation

A study on robot-assisted French pronunciation in a learning game

Liisa Peura
Doctoral researcher, French Department, School of Languages and Translation Studies, University of Turku
liisa.peura@utu.fi

Maarit Mutta
Adjunct professor, French Department, School of Languages and Translation Studies, University of Turku

Marjut Johansson
Professor, French Department, School of Languages and Translation Studies, University of Turku

Abstract
Robot-assisted language learning (RALL) is attracting growing attention among researchers and teachers thanks to its suggested positive impact on children's motivation and concentration in second language (L2) learning. However, the implementation in real classroom situations has not yet been widely investigated. Addressing this gap, this study explores the learning of French pronunciation with the help of a social robot as a part of the normal L2 classroom routine without the presence of a human teacher. A social robot has a capacity to interact autonomously or semi-autonomously with humans, and it has a physical embodied presence, unlike chatbots. Fifty Finnish primary school children, aged 10 to 12 years, participated in the pedagogical experiment, presented as a challenge game called *Défi Domi* that was a three-level game where learners were challenged by a set of pronunciation tasks. They learned by imitating the robot without explicit pronunciation rules. The design enabled exploration of the robot's role in learning pronunciation and its effects upon learners' persistence. The results showed that learners' pronunciation improved during the experiment and that pupils' active participation became a natural part of the learning process. The participants were persistent, and they enjoyed learning with the robot throughout the challenge. This study argues that gamified RALL designs can support interactive and autonomous learning when implemented in classrooms.

Keywords
Child robot interaction, robot-assisted language learning, L2 pronunciation, persistence

Introduction

Based on previous studies of different fields, it can be said that interaction is the medium of all learning (see e.g., Lee & Lee, 2022). Interaction is emphasized in language teaching and learning contexts—both in the classroom (formally) and outside the classroom (informally)—as learning to use a foreign language has inherently communicative and social goals (Eshach, 2007; Sundqvist, 2016). Thus, the recent development of robot-assisted language learning (RALL) has the potential to change the dimensions of language teaching. Robot-assisted language learning (RALL) is defined as "the use of robots to teach people language expression or comprehension skills—such as speaking, writing, reading, or listening" (Randall, 2019. p. 7:2). A meta-analysis by Lee and Lee (2022) suggests that RALL can have some superior pedagogical effects compared to non-RALL conditions, and social robots might therefore replace other technologies (a social robot has a capacity to interact autonomously or semi-autonomously with humans, and it has a physical embodied presence, unlike chatbots), although their potential to replace human teachers or peers is less likely.

RALL can bring improvements to language classrooms in several ways, and a child's eagerness to engage with robots, and the tendency to treat them as social beings, make child–robot interaction (CRI) a particularly fruitful approach to language learning (Ahtinen & Kaipainen, 2020; Baxter et al., 2017; Lee & Lee, 2022; Tanaka & Matsuzoe, 2012). Relating to learners' willingness to communicate (WTC), social robots have been found to reduce anxiety of speaking a second language (Belpaeme et al., 2018; Lee & Lee, 2022; Randall 2019). Language learning in a classroom usually occurs as one-to-many interactions. A robot tutor can prove helpful by allowing learners to engage in one-to-one language learning activities without the continuous, active involvement of the teacher (Belpaeme et al., 2015). Adaptive teaching programs with robots have the potential to provide the learner with individually tailored input, allocating focus to the learner's personal challenges (van den Berghe et al., 2019; Lee & Lee, 2022; Randall, 2019). Furthermore, robot technology can offer native-like language input (as it speaks with a native-like accent), accounting for teachers' and peer learners' diversity in linguistic skills (Belpaeme et al., 2015).

Social robots can be more advantageous than other forms of technological devices, since they are specifically designed for spoken interaction. The embodiment and physical presence, as well as the multimodal dimensions of a robot tutor, have been shown to enhance learning. Social robots can be more advantageous than other forms of technological devices, since they are specifically designed for spoken interaction (Ahtinen & Kaipainen, 2020; Belpaeme et al., 2015, Belpaeme et al., 2018; Han, 2012; Kennedy et al., 2016; Vogt et al., 2017). When communicating in a mother tongue (L1), learners can direct their attention more to the meaning of the message than to its form. Native-like acquisition is also suggested to be beneficial in a second language (L2) context, and it can be realized through meaningful interactions, such as dialogue and playing. This is easy to implement in RALL designs compared to traditional textbook-based learning. However, some studies claim that it is favorable to give the L1 a mediational role, helping to structure new knowledge about the L2 (Channa et al., 2017). To investigate the effectiveness of implicit learning through near native-like interactive engagement, the participants in this study interact in the target language with the robot, while their peer learners can support the process in their mother tongue.

The pedagogy of language learning and teaching has changed from an examination of learning outcomes to an examination of the learning process and active involvement (van Lier, 2010). According to the Common European Framework of Reference for Languages (CEFR 2018), pronunciation is intertwined in communicative language competencies and can be evaluated by three different scales: "overall phonological control," "sound articulation," and "prosodic features" (CEFR, 2018, pp. 134-135). To be efficient, interactive assignments must remain within the pupils' zone of proximal development (Vygotsky, 1978). When learning tasks are designed to be suitably challenging in relation to the learner's skills, they can promote optimal commitment and immersion in learning (Shernoff et al., 2009). In addition, when staying in the zone of proximal development, interactive games have the potential to evoke positive feelings and reduce the fear of failure, leaving learners to dare to make errors (Lee & Hammer, 2011). Studies have shown that when based on interaction, games support the development of oral language skills when the learner's entry level is low and the interaction partner is a native speaker (e.g., Canto et al., 2013). Including game-like elements in learning is suggested to increase motivation, even in boring tasks, further implying that improved motivation leads to better performance (Griggs et al., 2019).

While most recent studies have focused more extensively on young children's vocabulary learning in English in one-to-one interaction (Lee & Lee, 2022; Rintjema et al., 2018; van den Berghe et al., 2019), this study investigates the implicit French pronunciation learning of children aged 10 to 12 in a learning game with a social robot. Considering the influence of L1 on L2 learning (Yan, 2010), the present study provides unique insight into Finnish primary school French (L2) lessons. According to Randall (2019, pp. 7–12), there are only a few studies in RALL that concentrate directly on oral competence, such as pronunciation and its improvement in the classroom context (for e.g. Wang et al., 2009). Therefore, the present study is conducted through the following research questions:

1. What is the impact of a social robot on the learning outcomes of French pronunciation among primary school learners?
2. How does the robot motivate and enhance persistence in learning French pronunciation?

Previous findings

Potential benefits of RALL

Previous studies among children have found increased focus and motivation in RALL (e.g., van den Berghe et al., 2019; Vogt et al., 2017). Social robots have the ability to engage in eye contact and display nonverbal behavior, which can establish a persuasive and motivating space for interaction. For instance, the robot's simple nonverbal rewarding, such as multi-color blinking eyes, is found to be effective as a motivating element in terms of rehearsal and repetition (Ahtinen & Kaipainen, 2020). Findings on the communicative behavior of a robot, also used in this study, suggest that personalized and animated feedback is more enjoyable (Baxter et al., 2017). However, it is not clear whether robot feedback is crucial for the learning outcome (de Haas et al., 2016), and the excessive liveliness of the robot might even become distracting (Kennedy et al., 2015; van den Berghe et al., 2019).

One-to-one interaction with a robot has been suggested as leading to better learning outcomes than interacting with a robot in a group (Lee & Lee, 2022). A robot has been found to work best when it is introduced to learners as a friend or peer (Baxter et al., 2017; Tanaka & Matsuzoe, 2012). Research conducted on robotic companions in actual classroom settings suggests that robots show potential as effective tools also in relation to learning objectives. Even simple or repetitive learning tasks can become more interesting (Castellano et al., 2013). This positive impact is partly speculated (van den Berghe et al., 2019) to be due to the enthusiasm generated from using a new tool. Similarly, Randall (2019) argues that there is a need for longitudinal studies in real-world settings to study the learning effects after initial exposure. Based on the recent meta-analysis by Lee and Lee (2022), one can argue that it is not about the novelty effect per se. RALL increases motivation, and more longitudinal studies are needed to study this effect in more detail.

Teaching pronunciation

In Finland, due to the emphasis on communicative language teaching, pronunciation has not received significant attention in L2 teaching, with the focus being on effectively conveying the message (Tergujeff, 2013). Nowadays, clear and comprehensible pronunciation is considered one of the most crucial objectives for pronunciation learning, rather than aiming for native-like pronunciation (Tergujeff & Kautonen, 2017, p.18). It is vital for learners to understand the significance of pronunciation in terms of achieving comprehensibility.

According to Jügler and Möbius (2015), corrective feedback is needed in both perception and production, and it should focus on those factors where learners face the most challenges. The phonetic and phonological differences of the learner's L1 and L2, both at the sound (i.e., segmental) and prosodic level, might influence learning (Jügler & Möbius, 2015).

Together with the teacher's feedback, the learner's self-evaluation is also an important factor in learning pronunciation. Listening to one's own pronunciation in comparison with the model helps one notice one's own mistakes. However, this can be challenging, because not all learners hear the difference between their version and the correct model (Couper, 2006). Thus, RALL could support teachers by offering new resources and a more varied toolkit to teach pronunciation, providing authentic input. One major technical concern is that speech recognition systems are not sufficiently developed to support child-robot interaction. As pointed out by Belpaeme et al. (2015), the characteristics of a child's speech might differ significantly from adult speech, which sets special requirements for natural language processing systems in social robots. Limited speech recognition also brings challenges to the teaching space (Kennedy et al., 2017). To bridge this gap in research, this study explores the robot's suitability in pronunciation learning when the robot is left to run semi-autonomously without being controlled by the experimenter.

Methods

Presentation of the study

This study explores the functionality of RALL in an experiment in which primary school pupils learn words in dyads, using a simple repetition learning game challenge, without explicit pronunciation rules, and without the presence of a human teacher. The three-level challenge game called *Défi Domi* was designed to provide a progressively more difficult set of pronunciation tasks/tongue twisters for the learner to overcome. The understanding of the target words was facilitated by pictures to keep the challenge game purely in French. The explicit Finnish translations of the learning targets were not used in this experiment, to encourage implicit learning.

As Fryer and Carpenter (2006) showed in their studies, learners benefit from the fact that chatbots do not lose their patience while repeating the same content several times. Furthermore, Bushweller (2020) suggests that the strength of the robot is its ability to provide learning practice regardless of place and time without fatigue. Based on this, together with the physical interaction that RALL methods allow, it is hypothesized that the robot is effective in teaching children L2 pronunciation: children learn the right pronunciation from the robot (H1). Based on the aforementioned theories and guidelines for recommended practice presented by Robinson et al. (2019), the hypothesis (H2) is that a robot-assisted game-like setting engages learners in a positive way. Learners are focused, and motivated to engage in the game, without paying attention to the number of repetitions.

Data collection and participants

The pedagogical experiment was carried out in a Finnish primary school with pupils aged 10 to 12 years (fourth grade and sixth grade). All of the children started their French studies in the 3rd grade at the age of nine, and French was their first foreign language. With school and parental consent, a total of 50 pupils participated in the experiment: 24 sixth-grade students (8M, 16F) and 26 fourth-grade students (7M, 19F). The participants were made familiar with the robot before the study began. The robot was presented as a new mascot pupil at the beginning of the semester, and was called Dominique, which is a gender-neutral name in French.

The data were collected through tests and questionnaires. Before the challenge game experiment, the starting skill level was recorded and scored to measure the progress of learning after the sessions with the robot. The participants saw *Défi Domi*'s challenge words (see section 3.3 and Appendix 1) for the first time in the beginning of the recording session, and they read them out loud for the research assistant before starting the game session with the robot. Thus, the first recordings were not affected by the robot. After the challenge game sessions with the robot sessions, new recordings were made – the participants read out loud the same wordlists. These recordings were used to analyze the eventual learning of pronunciation. To analyze the progress in pronunciation, the recorded reading of the word lists was rated by the experimenter and a second rater both before and after the experiment, to make the scoring more objective. In the case of differing scores, the two raters agreed on a final score. This approach is meant to increase the validity and reliability of the results, because of inherent subjectivity in auditory analysis.

In addition, the experiment included anonymous online questionnaires, both before and after the learning sessions. All participants answered multiple choice questionnaires which included statements on the scale 1-6. The pre-questionnaire covered the learners' general attitude towards learning French and French pronunciation—what they thought was the most difficult part of learning French and what they enjoyed the most. The post-questionnaire focused on the learning experience and the effects—emotions evoked during the sessions, self-evaluation of the learning session, and the efficiency of the learning method (see Figures 4 and 5). Both questionnaires also included three optional open-ended questions: "Briefly describe how you feel about the robot" (49 answers), "I would prefer to learn by other means" (24 answers), and "Other comments" (24 answers). These answers, paired with the experimenter's observations, were used for the qualitative analysis of the study.

The robot and the challenge game

The challenge game was carried out using a SoftBank Robotics Nao 6 robot (57.4 cm tall), standing/sitting in front of the participant and the browser-based application of the appropriate Elias software.

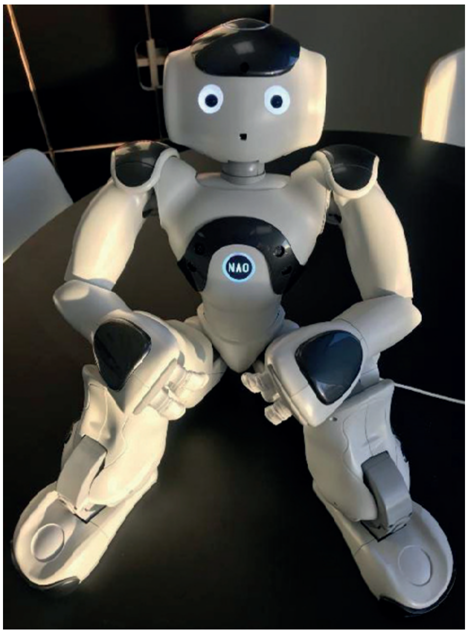

Figure 1
SoftBank Robotics Nao 6 robot, called "Dominique".

During the two-week experiment, the robot taught the participants a set of words, sentences, and tongue twisters. The experiment was presented as a challenge game called *Défi Domi* with three different levels: Level 1 consisted of single words with nasal sounds (e.g. *imprimer* '*to print*', *un enfant* '*a child*'); Level 2, consisted of words with fricative sounds /ɛ̃/, /œ̃/, /ɔ̃/; /z/, /ʃ/, /ʒ/ and silent letters (e.g. *un cochon rose* '*a pink pig*', *une bougie jaune* '*a yellow candle*'); and Level 3 consisted of challenging tongue twisters combining the two previous target sounds (e.g. *cent onze tranches de jambon blanc* '*one hundred and eleven slices of white ham*'). The researchers chose the set of ten words and sentences at each level, based on the school's French teachers' experiences with the most difficult challenges facing the pupils in terms of phonetics and pronunciation in French. French is not a common language heard in Finland. It was therefore assumed that the children were not exposed to the target words outside the experimental sessions.

In the task, learners sought to reach and complete the most difficult level. The robot pronounced a target word/sentence/ tongue twister and then the participant repeated the sentence. The robot did not continue to the next sentence before the participant pronounced the target sentence correctly (i.e., the speech recognition system recognized it). The recognition threshold was set to 40 %. The learning game was made persuasive by "candy eyes," which means the robot's eyes lit up colorfully with a sound effect as a reward for correct pronunciation. Understanding the meaning of the target words was not essential, but the participants were able to see the word as an image on the screen (e.g., *a pink pig*). The learning of the correct pronunciation was not based on any explicit rules except the simple imitation of the robot, which then validated the correct pronunciation.

Procedure

The pupils were randomly placed into pairs or teams of three. Each team had two weeks to work with the robot and advance in the challenge game at their own pace. They were able to use the French lessons, breaks, or time given by the class teacher to participate in the challenge. Thus, the individual time spent on the task varied depending on the motivation of the participant. The experimenter was present in the classroom next door in case of problems with the experiment that could prevent continuation of the exercise (i.e., errors in the robot's code).

While one learner performed the task with the robot, the other used a tablet to control the progress in the learning task and recorded the number of repetitions on a chart designed for the purpose. On the chart, the peer could see the words in French as well. In addition, some remarks on the performance and comments on the robot were listed by the controlling peer (e.g., *you did really great, bugs – the robot fell off line, did not work properly*). The written charts allowed the teacher to track the progress during the two weeks experimental period. The number of repetitions of the learners was used as a measure of perseverance in the study. This data was included for each learner for all the target words they reached and practiced. Practically all pairs managed to reach Level 3 within the two weeks, although not all of them succeeded in finishing it. After the gaming sessions, new pronunciation recordings were made, and the learners read the same word lists used at the beginning of the study. These recordings were used to analyze eventual learning in pronunciation.

Data pre-processing and statistical analysis

At each challenge level, every target was scored using a scale between 0 and 1. Level 1 consisted of single words (either verbs or nouns with articles), and the pronunciation of the target sound was scored with either 0 or 1. The scoring was strict. If there was a single error in pronunciation (including an article error), no point was given. Level 2 and 3 consisted of phrases, in which 0.5 points were given for semi-correct pronunciation wherein the pupil managed to successfully pronounce the main sounds of the phrase (for further details, see Appendix 1).

To explore Hypothesis 1, the impact on the learning of correct pronunciation, a statistical analysis of the pronunciation tests was conducted in RStudio (RStudio team, 2019). A paired one-sided t-test was conducted across all participants' ratings to see whether the scores were statistically greater in the posttest compared to the pre-test. To analyze the relationship between the number of repetitions and learning outcomes, a Pearson's correlation test was conducted. Finally, the Pearson's correlation test was used to measure the linear relationship between numeric answers to questions measuring motivation and the number of repetitions.

Results

Learning outcomes

The results of the statistical analyses suggest that the pedagogical design of robot-assisted learning is beneficial for acquiring a second language. The statistical results show learning development at all levels of the challenge game for both fourth and sixth graders.

Learning outcomes

Grade	*Game level*	*Mean improvement*	*SD*	*Mean repetitions*	*SD of repetitions*
4	1	2.52	1.34	21.89	11.2
4	2	2.1	1.64	56.29	37.11
4	3	2.41	1.13	106	61.3
6	1	2.44	2.12	14.78	3.61
6	2	1.8	1.32	24.08	14.9
6	3	2.25	1.3	46.67	26.01

Table 1
Mean improvement and number of repetitions per grade and level

At all levels, the descriptive statistics from Table 1 show the mean improvement and number of repetitions by grade and game level, with corresponding standard deviations, indicating that pronunciation was improved. This was supported by the statistical analysis: the t-test revealed that the post-scores ($M = 6.77$, $SD = 1.47$) in pronunciation were significantly higher than the pre-scores ($M = 4.72$, $SD = 1.64$); $t(49) = 16.21$, $p < .001$. Figure 2 illustrates the results of the t-test for all three levels of the challenge game with all participants.

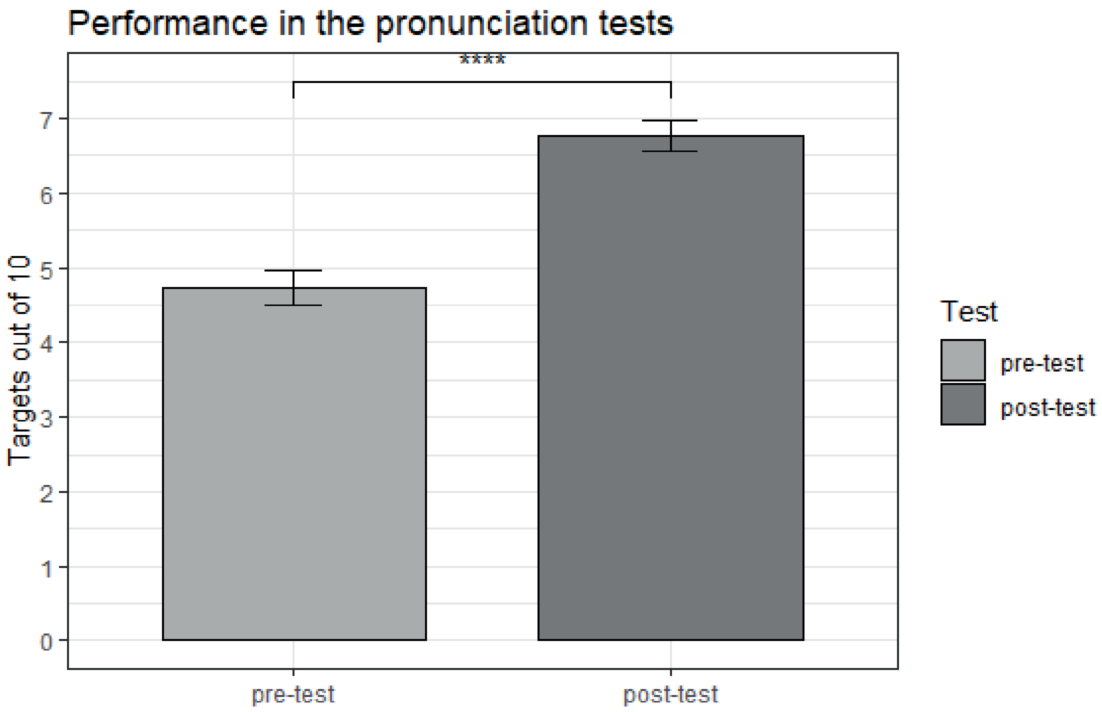

Figure 2
A visualization of the t-test with SE.

The small standard errors of the mean indicated that the learning was truly beneficial for the participants. These results support H1.

Persistence

A Pearson's correlation test was conducted between the number of repetitions and the participant's answers to the motivation questionnaire (on an agreement scale from 1 to 6). The test was not significant, implying that even a high number of repetitions did not decrease motivation during the challenge. This is supportive for H2. The repetitive nature of the task was still experienced as engaging.

All learners sought to reach Level 3 and correctly pronounce the difficult sounds, which required persistence, especially considering the encountered speech recognition deficiencies. The number of repetitions per learner was used as a measure of persistence (shown in Table 1), and the preliminary results emphasized the perseverance of the learners. To explore Hypothesis 2, the correlation of the robot on persistence in learning correct pronunciation was viewed in light of repetitions in relation to the learners' learning experience. It was expected that the robot would have a positive association on persistence, and the results positively correlated with this expectation. It is important to note the substantial number of repetitions recorded in the data. Despite the robot's intermittent speech recognition problems, learners periodically repeated targets up to several dozen times. In general, the learners persevered, and only a few gave up at Level 3. *Défi Domi* was a challenge game, but it was not designed to be a competition between groups. However, the younger pupils liked to think of the challenge game as a race. This might explain the eagerness to proceed and the high number of repetitions. The large number of repetitions implies that the willingness to communicate and try different sounds encouraged perseverance when interacting with the robot.

Remarks on affective states

To study how the robot enhances and maintains learners' motivation (cf. Dörnyei, 2001) during and after the CRI, we studied learners' experienced emotions and affective states through exploratory questionnaires. When performing the challenge task, the learners were independent, and active learners worked with the robot in pairs without the help of a teacher. The positive association of the social robot on pronunciation learning was explored through the participants' experiences. First, they were asked to rate four statements about self-efficacy on a scale from one to six (1 = totally disagree to 6 = totally agree), as shown in Figure 4.

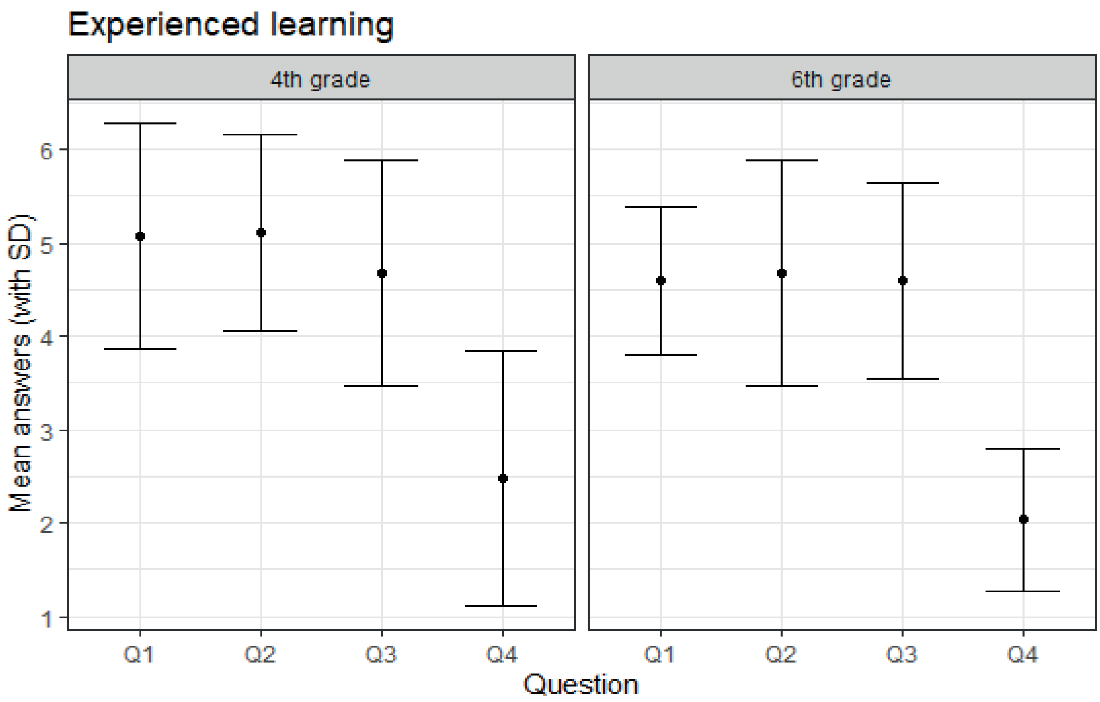

Figure 4
Average answers of the experienced learning. Q1 = I learned to pronounce French better, Q2 = I realized new rules about French pronunciation, Q3 = I would be able to use the studied sounds on my own, Q4 = I think I would have learned better using another method.

Overall, the high scores for the first three questions (Q1–Q3) indicate that the learners regarded the RALL learning game as beneficial for learning, and rated the activity highly in terms of their own competence. The low score on Q4 suggests that the method was liked. Nevertheless, to motivate their learning, learners should feel challenged.

The results (Figure 5) imply that the learning experience was positive and engaging. The participants were given the following statements (Figure 5), and asked to answer on a scale from 1 to 6 (totally disagree to totally agree).

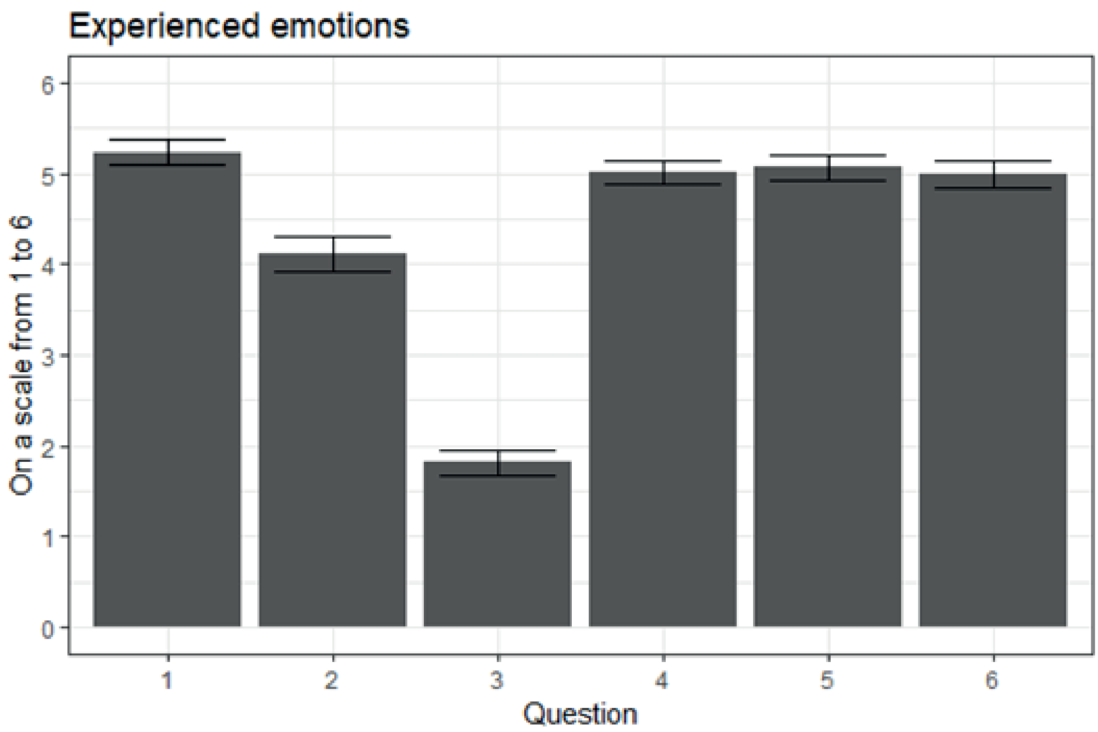

Figure 5
Experienced emotions by the participants themselves. Q1: I was excited, Q2: I was slightly nervous, Q3: I was anxious, Q4: I was focused, Q5: I experienced joy of learning, Q6: The exercise put me in a good mood.

The results indicate that working with a robot produces joy in the learning process. Although the learners were initially nervous about talking to the robot in this experiment, the exercise put them in a good mood. There was little difference in the level of emotion experienced between the different age groups. The role of affective states is worth noticing: despite the simple nature of the repeating task and the robot's intermittent speech recognition problems, RALL appeared to be fun and useful. The participants had positive attitudes towards RALL; emotions of anxiety were minimized.

The majority of the participants found it easier to talk to the robot, since the robot did not judge them personally, as can be seen in the following open-ended responses of the post-test questionnaire (originally in Finnish):

> "It's wonderfully relieving when I don't have to think about what Domi thinks of me."
>
> "It's easy to talk to because it cannot criticize."
>
> "It's nicer to talk to Domi than in class with other learners."

Limitations of this study

The core curriculum for basic education (FNBE, 2016) regulates all teaching in Finland. As this experiment was carried out in a classroom setting as a part of ordinary French teaching, the design of the experiment was carefully considered to serve educational goals along the tracks of the curriculum. Therefore, the setting was designed in a game format to be played in pairs thus promoting intrinsically social and interactive competences. Furthermore, according to the CEFR (2018), in a modern classroom teaching methods are learner-centered, and consequently, the learning method in this experiment was tailored for the participants: the target challenges were based on the teacher's assessment of the most difficult sounds for the group. We used descriptive statistics to support our hypotheses, which allowed for exploration of whether the method worked with participants of two age groups with different pronunciation difficulties, and it was also designed to guide future teaching practices. Using a control group outside of these participants would have given it more credibility, however, it would also have made the comparison invalid, due to uncontrollable variables, such as different teaching materials, a different teacher of participants, and potentially differing pronunciation challenges. All the same, in future studies, a counterfactual analysis with a control group would deliver more robust results for verifying the impact of the pedagogical intervention.

Discussion

The study explored the robot's role in learning pronunciation, learning outcomes, and the interactive experience with a social robot without the presence of a teacher. Learning outcomes supported H1, since the results indicated the pronunciation of the participants was improved through the interactive learning game. Our findings suggest that pronunciation can be learned through repetition (cf. Tergujeff, 2013) in interaction in dyads with a robot speaking only the target language. However, as pointed out by Channa et al. (2017), L1 can play a mediational role in L2 learning, which was also the case in this experiment: the learners used Finnish when working on the targets with their partner, structuring their knowledge of the L2. Moreover, the current design nudged pupils to assist each other, and incurred both implicit learning through repetition and explicit learning through pair work. The peer learner who was recording the progress on the chart often assumed the teacher's role: they heard the mispronounced part better than the learner, and pointed out the challenging sound (i.e. negotiation). In line with Jügler and Möbius (2015), self/peer assessment facilitated learning.

The peer's support also helped in situations where the limitations of the robot's speech recognition system brought challenges. The failed communication with the robot led to many problem-solving situations among the participants, increasing language awareness as they tutored each other and provided corrective feedback. As observed, the interaction flowed naturally without the teacher's input. Learning with others is found to lead to better outcomes than individual-centered models (Kyndt et al., 2013). However, the risk of frustration exists, due to limited language processing systems in social robots, as pointed out by Belpaeme et al. (2015). Thus, in contrast to Lee and Lee's (2022) suggestion on one-to-one interaction with a robot, the dyad collaboration might have helped both learning correct pronunciation and managing frustration, although its effect on learning was beyond the scope of this study.

As for motivation and persistence (H2), the high number of repetitions indicated that the pupils were persistent in doing the task, and the statistical results support this. The results show the repetitive nature of the task did not decrease their motivation during the experiment. In line with Robinson et al. (2019), the results of the post-test questionnaire further suggest that a robot-assisted, game-like interaction setting enhances long-term engagement. The large number of repetitions implies that the willingness to communicate encouraged perseverance when interacting with the robot. The interaction with the dyad and the robot created a safe space for proximal development (Vygotsky, 1978), where the participants could try out new sounds without being afraid of mistakes.

In natural interactions, the benevolent listener already understands half of the speaker's intention. Since the robot does not have the capacity to understand partially correct responses, it results in more repetition for the learner than would be required in real-world situations. Ultimately, the joy of success with the robot appeared to be greater than the technological problems for the duration of the experiment. The motivating factor of the robot seemed to be its somewhat clumsy multimodality, which helped engender a sense of friendliness in the interactive experience. The participants' responses supported the assumption that a robot reduces anxiety about speaking a foreign language (Belpaeme et al., 2018; Lee & Lee, 2022; Randall 2019). As Lee and Lee (2022) pointed out in their meta-analysis, more research is needed on acquiring oral competence in RALL. The results of this study support the use of the robot also for strengthening oral skills in terms of pronunciation.

Related to interaction dynamics, this novel application of embodied and gamified methods in L2 learning with the assistance of a social robot also changes the role of the teacher. In RALL, the role of the teacher as the moderator of the conversation is diminished, but there is a new emphasis on designing and creating a learning task coded for the robot that enables scaffolding (Vogt et al., 2017), as designed in the present experiment.

Conclusion and future implications

The aim of the study was to investigate the facilitation of learning a second language using RALL with an applied game design. The experiment was designed on the basis of general pronunciation challenges in acquiring a new language, with an exploratory focus. The quantitative results were accompanied by qualitative observations that identified potential areas for future research, such as long-lasting effects on pronunciation and collaboration between learners.

The results imply that the robot is perceived as a motivating factor producing improved learning in pronunciation, even without the novelty effect of the robot. The qualitative observations indicate the target sounds in French that do not exist in Finnish were learned in a positive way. In line with previous findings (Ahtinen & Kaipainen, 2020), active participation became a natural part of the learning process with the robot. Without emphasis, the peer learners became aware of their teammate's pronunciation challenges, implicitly adopting French pronunciation tendencies. It should be noted that the repetitions of targets may have induced implicit vocabulary learning, which extended beyond the scope of this experiment.

Gamified RALL produced improved self-efficacy, although technical limitations remain a barrier to efficient implementation and natural interaction. In the future, a more adaptive robot tutor could adjust to the level of each learner to focus on individual challenges in an interactive and safe environment. With an improved speech recognition system and adaptive tutoring, the robot could enhance this effect of language awareness and allocate even more focus to each individual's learning challenges. Therefore, future studies could focus on the role and the degree of sophistication of the robot feedback, taking into account the restrictions on data protection in a primary school context.

Our study indicates that two learners can engage together in a learning task, creating a shared interaction space with a robot. It would be interesting to compare the current dyad-robot design with a single child-robot design, and a child-child design, to reveal the benefits of each design – both for motivation and learning outcomes and the nature of the interaction space. The core aspects of the present setting, the robot's physical presence, learning pronunciation through interaction with a peer, and self-evaluation, require active engagement, and thus the intrinsically embodied and social nature of the task might even have a positive effect on learners' empathy skills. Thus, dyad interaction and the role of the dyad's collaborative reasoning to solve problems related to pronunciation and robot interaction could be potential areas for future research. Nevertheless, robots can currently play a role in enriching teaching practices: learning pronunciation, motivating learning, and enhancing interaction.

Acknowledgements. We are grateful to the project researcher Anni Nieminen for her work in the data collection and Telma Peura for her help with R programming. We thank participating children and their parents for their assistance in this research as well as the city of Espoo, Utelias Technologies Oy and Elias the robot, "Dominique". We are particularly thankful for the reviewers for their valuable comments. We also thank Finnish Cultural Foundation's Varsinais-Suomi Regional Fund for supporting this work.

References

Ahtinen, A., & Kaipainen, K. (2020). Learning and Teaching Experiences with a Persuasive Social Robot in Primary School – Findings and Implications from a 4-Month Field Study. In S.B. Gram-Hansen, T. S. Jonasen, & C. Midden (Eds.), *Persuasive Technology. Designing for Future Change* (pp. 73–84). Springer International Publishing. https://doi.org/10.1007/978-3-030-45712-9_6

Baxter, P., Ashurst, E., Read, R., Kennedy, J., & Belpaeme, T. (2017). Robot education peers in a situated primary school study: Personalisation promotes child learning. *PLOS ONE*, *12*(5), e0178126. https://doi.org/10.1371/journal.pone.0178126

Belpaeme, T., Kennedy, J., Baxter, P., Vogt, P., Krahmer, E. E. J., Kopp, S., Bergmann, K., Leseman, P., Küntay, A. C., Göksun, T., Pandey, A. K., Gelin, R., Koudelkova, P., & Deblieck, T. (2015). L2TOR – Second Language Tutoring using Social Robots. *Proceedings of the ICSR* 2015 WONDER Workshop. Springer.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, *3*(21).https://doi.org/10.1126/scirobotics.aat5954

Bushweller, K. (2020). Teachers, the robots are coming. But that's not a bad thing. *Education Week*,*7*. https://www.edweek.org/technology/teachers-the-robots-are-coming-but-thats-not-a-bad-thing/2020/01

Canto, S., Jauregi, K., & Bergh, H. van den. (2013). Integrating cross-cultural interaction through video-communication and virtual worlds in foreign language teaching programs: Isthere an added value? *ReCALL*, *25*(1), 105–121. https://doi.org/10.1017/S0958344012000274

Castellano, G., Paiva, A., Kappas, A., Aylett, R., Hastie, H., Barendregt, W., Nabais, F., & Bull, S. (2013). Towards Empathic Virtual and Robotic Tutors. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Artificial Intelligence in Education* (pp. 733–736). Springer. https://doi.org/10.1007/978-3-642-39112-5_100

CEFR 2018 Council of Europe. Common European framework of reference for languages. Learning, teaching, assessment. Companion volume with new descriptors. Cambridge University Press. www.coe.int/lang-cefr

Channa, L. A., Gilhooly, D., Lynn, C. A., Manan, S. A., & Soomro, N. H. (2017). Friend or foe? First language (L1) in second/foreign language (L2/FL) instruction & Vygotsky. *Journal of Language and Cultural Education*, *5*(2), 25–39. https://doi.org/10.1515/jolace-2017-0016

Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect: an Australian Journal of TESOL*, *21*(1), 46–66. http://hdl.handle.net/1959.14/329829 de Haas, M., Vogt, P., & Krahmer, E. (2016). Enhancing child-robot tutoring interactions with appropriate feedback. In Proceedings of the Long-Term Child-Robot Interaction Workshop, at RO-MAN 2016

Dörnyei, Z. (2001). Maintaining and protecting motivation. In *Motivational Strategies in the Language Classroom* (Cambridge Language Teaching Library, pp. 71-116). Cambridge University Press. https://doi.org/10.1017/CBO9780511667343.005

Eshach, H. (2007). Bridging In-school and Out-of-school Learning: Formal, Non-Formal, and Informal Education. *Journal of Science Education and Technology*, *16*(2), 171–190. https://doi.org/10.1007/s10956-006-9027-1

FNBE. (2016). National core curriculum for basic education 2014. Finnish National Agency for Education.https://www.oph.fi/en/statistics-and-publications/publications/new-national-core-curriculum-basic-education-focus-school

Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, *10*(3), 8–14. http://dx.doi.org/10125/44068

Griggs, A., Lazzara, E. H., Palmer, E., Fouquet, S., Leverenz, T., Raushel, A., & Doherty, S. (2019). Utilizing Games for Learning: Applications of Game-Based Training and Gamification. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 2166–2168. https://doi.org/10.1177/1071181319631361

Han, J. (2012). Robot assisted language learning. *Language Learning & Technology*, *16*(3), 1-9. http://dx.doi.org/10125/44291

Jügler, J., & Möbius, B. (2015). Auditory feedback methods to improve the pronunciation of stops by German learners of French. In M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. Scobbie (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow, UK. 10-15 Aug, 2015* (p. 5). International Phonetic Association.

Kennedy, J., Baxter, P., & Belpaeme, T. (2015). The Robot Who Tried Too Hard: Social Behaviour of a Robot Tutor Can Negatively Affect Child Learning. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 67–74. https://doi.org/10.1145/2696454.2696457

Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2016). Social robot tutoring for child second language learning. *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 231–238. https://doi.org/10.1109/HRI.2016.7451757

Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F., Senft, E., & Belpaeme, T. (2017). Child Speech Recognition in Human-Robot Interaction: Evaluations and Recommendations. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 82–90. https://doi.org/10.1145/2909824.3020229

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational research review*, *10*, 133-149. https://doi.org/10.1016/j.edurev.2013.02.002

Lee, J., & Hammer, J. (2011). Gamification in Education: What, How, Why Bother? *Academic Exchange Quarterly*, *15*(2), 1–5.

Lee, H., & Lee, J. H. (2022). The effects of robot-assisted language learning: A meta-analysis. *Educational Research Review*, *35*, 100425. https://doi.org/10.1016/j.edurev.2021.100425

Randall, N. (2020). A Survey of Robot-Assisted Language Learning (RALL). *ACM Transactions on Human-Robot Interaction*, *9*(1), 1–36. https://doi.org/10.1145/3345506

Rintjema, E., van den Berghe, R., Kessels, A., De Wit, J., & Vogt, P. (2018). A robot teaching young children a second language: the effect of multiple interactions on engagement and performance. *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 219-220. https://doi.org/10.1145/3173386.3177059

Robinson, N. L., Turkay, S., Cooper, L. A. N., & Johnson, D. (2019). Social Robots with Gamification Principles to Increase Long-Term User Interaction. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 359–363. https://doi.org/10.1145/3369457.3369494

RStudio Team. (2019). RStudio: Integrated Development Environment for R. Boston, MA.

Shernoff, D. J., & Csikszentmihalyi, M. (2009). Cultivating engaged learners and optimal learning environments. *Handbook of positive psychology in schools*, (pp. 131–145). Routledge.

Sundqvist, P. (2016). Gaming and young language learners. In F. Farr & L. Murray (Eds.), *The Routledge Handbook of Language Learning and Technology* (pp. 446-458). Taylor & Francis.

Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, *1*(1), 78–95. https://doi.org/10.5898/JHRI.1.1.Tanaka

Tergujeff, E. (2013). English pronunciation teaching in Finland. *Jyväskylä Studies in Humanities*, *207*. https://jyx.jyu.fi/handle/123456789/41900

Tergujeff, E. & Kautonen, M. (2017). Suullinen kielitaito ja sen tärkeys. [Spoken competency in its importance] In E. Tergujeff & M. Kautonen (Eds.), *Suullinen kielitaito: Opi, opeta, arvioi* [Spoken competency: Learn, teach, evaluate] (pp. 12–21). Helsinki: Otava.

van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., Van der Ven, S., & Leseman, P. (2019). Social robots for language learning: A review. *Review of Educational Research*, *89*(2), 259-295.

van Lier, L. (2010). The ecology of language learning: Practice to theory, theory to practice. *Procedia – Social and Behavioral Sciences*, *3*, 2–6. https://doi.org/10.1016/j.sbspro.2010.07.005

Vogt, P., de Haas, M., de Jong, C., Baxter, P., & Krahmer, E. (2017). Child-Robot Interactions for Second Language Tutoring to Preschool Children. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00073

Vygotsky, L. S. (1978). Mind in Society: Development of Higher Psychological Processes. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Wang, Y. H., Young, S. S., & Jang, J. S. R. (2009). Evaluation of tangible learning companion/robot for English language learning. *2009 Ninth IEEE International Conference on Advanced Learning Technologies* (pp. 322-326). IEEE. https://doi.org/10.1109/ICALT.2009.147

Yan, H. (2010). The Role of L1 Transfer on L2 and Pedagogical Implications. *Canadian Social Science*, *6*(3), 97–103. https://doi.org/10.3968/j.css.1923669720100603.012

Appendix 1.

Pronunciation description and the criteria used for scoring (the most typical errors with 0 points):

LEVEL 1	1 p	0,5 p	0 p
le français	[ləfʀɑ̃sɛ]		[lefʀɑ̃sɛ] an error in the pronunciation of the article /le/ instead of /lə/.
imprimer	[ɛ̃pʀime]		[impʀime] [ɛ̃pʀimerʀ] nasal sound /ɛ̃/ was missing or the /ʀ/ sound was heard in the end.
LEVEL 2	1 p	0,5 p	0 p
un cochon rose	[œ̃kɔʃɔʀoz]	[œ̃ʃɔʃɔ̃ʀoz] [œ̃kɔʃɔ̃ʀos] words were otherwise correctly pronounced but there was one significant phonetic error: in the beginning of the word the letter c was pronounced /ʃ/ not as /k/ as it should, or if the s in the word *rose* was pronounced /s/ and not as not /z/ as it should.	[œ̃ʃɔʃɔ̃ʀo:s] the four key sounds were missing: ~/ʃɔʃɔro:s/, or there were two phonetic errors in the main words and the article was mispronounced (the nasal sound was missing).
j'ai une girafe	[ʒeynʒiʀaf]		[ʒaiungiʀaf] the key phonemes were missing and the article was mispronounced: /ʒaiungiraf/
LEVEL 3	1 p	0,5 p	0 p
Charles joue à chat chaque jour avec Jean	[ʃaʁlʒuaʃaʃakʒuravɛkɑ̃]	[ʃaʁlʒuɔʃaʃakueʒuravɛkʒɑ̃] the preposition *à* was pronounced as " *au* " and the word *chaque* was mispronounced, but the s-sounds were correct.	[ʃaʁlʒuɔʃatʃakueʒuavɛkʒœn] the /t/ was pronounced in the end of the word *chat*, and the letters *ue* were pronounced in the word *chaque*, if the /ʀ/ sound was missing in the word *jour* and if *Jean* was pronounced as *jeune*.
Ton doudou est tout doux	[tɔ̃duduɛtudu]	[tɔ̃duduɛtuduz] [tɔ̃duduɛtudus] the diphthong /ou/ was correctly pronounced, but the x in the word *doux* was pronounced as /z/ , or /s/ was heard in the word.	[tɔ̃doudouɛstoudouz] [tɔ̃doudouɛstoudous] if the phrase was pronounced as in Finnish /tou/ and /dou/, and if the /s/ was heard in the word *est.*