INTERSPEECH 2021: Show & Tell Contribution
30 August – 3 September, 2021, Brno, Czechia

Lalilo: a reading assistant for children featuring speech recognition-based reading mistake detection

Corentin Hembise[1], *Lucile Gelin*[1,2], *Morgane Daniel*[1]

[1]Lalilo, France
[2]IRIT, Paul Sabatier University, CNRS, Toulouse, France
{corentin, lucile, morgane}@lalilo.com

Abstract

Lalilo is a reading assistant intended to help kindergarten to second grade students to master their reading skills. Students progress at their own pace thanks to an adaptive learning system that differentiates instructions. Teachers can access data on their students' progression. Among other exercises, a read-aloud exercise is provided for students to practice their reading. This exercise uses a reading mistake detection system based on speech recognition to offer automatic feedback on the child's reading. Since speech recognition on children learning to read is highly challenging, we overcome potential inaccurate thus damageable feedback with an uncertainty estimation leading to a neutral feedback.

Index Terms: reading mistakes detection, child speech, reading assistant, uncertainty estimation

1. Introduction

Learning to read is a fundamental step in the process of becoming an informed citizen. Furthermore, it was shown that difficulty in reading can affect the whole intellectual development of children. To become proficient readers, it is essential for students to practice reading aloud. Oral practices have indeed a positive impact on word recognition, fluency, comprehension and improves reading achievement [1]. For oral practices to be efficient, it must be done regularly and include guidance and/or feedback [2]. Not all children have the support to practice read aloud at home and it is an important logistic in a classroom setting as teachers need to listen to students sequentially.

Several reading tutors have been developed [3, 4, 5] to help teachers in this tremendous task. Speech recognition for young children learning how to read is particularly challenging as it combines child speech well known acoustic and prosodic difficulties [6, 7] to reading miscues brought by non proficient readers. [8] confirms that the younger the children, the harder the task. Reading tutors are often used in classroom settings where babble noise can alter recordings' quality. In addition to an already challenging task, reading tutors cannot permit themselves to provide inaccurate feedback on student's reading, as it can be particularly damaging for the child's progression, and generate disengagement and frustration.

In this demonstration, we introduce the Lalilo platform: a reading assistant for 5 to 8 year-old students, empowered by a speech recognition based reading mistake detection enabling learners to practice reading aloud with immediate feedback. We present how we avoid giving wrong feedback to the students with a feedback regulation based on uncertainty estimation. Finally, the teacher interface allows them to access their student's recordings and the corresponding feedback.

2. Presentation of the platform

Lalilo developed a pedagogical assistant to help English and French native children learn how to read. It is developed for Kindergarten to second grades, i.e. children aged from 5 to 8 years. It is available through a web application via computers or tablets. It is designed to be mainly used at school in autonomy, as to offer teachers the possibility to do small groups differentiated instruction, or at home under parent supervision. The pedagogical assistant is designed to fit each student's needs: a personalized pathway is created throughout their progression. This personalized pathway is generated by an adaptive learning module, that chooses the best exercise for each student depending on their pedagogical progression and previous answers. The gamification of the application with a rewards system and different worlds to discover helps the students to keep engaged. Teachers can see student's progression through a dedicated dashboard and exploit the available data for differentiated instructions.

2.1. Read-aloud exercise

Among many various exercises, Lalilo offers a read-aloud practice through an exercise, on all devices with supported browsers and equipped with a microphone. The student is asked to read various content, according to the difficulty level chosen by the adaptive learning algorithm. The content varies from isolated units such as words, letters or syllables, to longer contents such as sentences or lists of units. The system predicts reading mistakes on one or several words from the content (the target words) and give feedback to the student.

As described in Fig.1, the learner gets instructed to read in a clear voice the written target text. Once they finish recording themselves, their recording is played back and they have the possibility to record themselves again if they are not satisfied. The recording and content are sent to the speech recognition API that computes and returns feedback on the learner's reading. The feedback can be either positive (when the reading is correct), encouraging (when a reading mistake has been detected) or neutral (when the system is uncertain). In the latter case, the system asks the student to listen the correct reading and self-evaluate them. The entire flow has been thought to be used by non-readers by reading them instructions and using icons instead of words.

2.2. Feedback computation

The computation of the feedback is a python API accessed by the front-end (the exerciser). Given a recording and the text to read, the system applies diverse speech recognition and classification modules that aim at detecting reading mistakes on the target word(s) [9].

As noted earlier, speech recognition for children is chal-

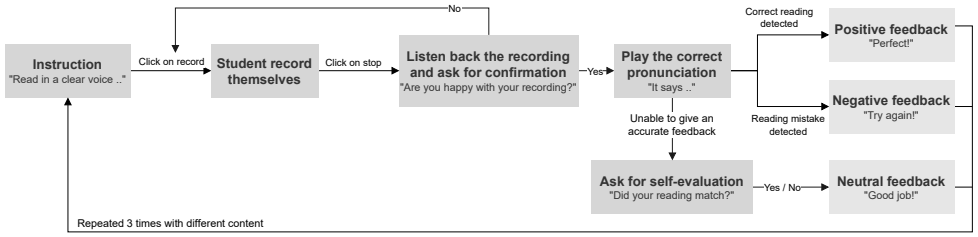

Figure 1: *Read-aloud exercise's flow.*

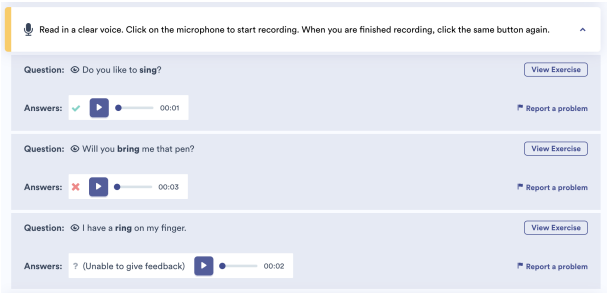

Figure 2: *Teacher's view on student's recordings.*

lenging. Giving an incorrect feedback to students can be frustrating and even damaging for their learning, in particular when the reading is correct and the feedback is negative. While the system gets continuously improved to make it robust to the particular classroom environment, mistakes are hardly avoidable. Hence, the system needs to work with them. An uncertainty brick based on the Trust Score algorithm [10] have been developed for this purpose. This module is build to quantify the uncertainty associated to the reading feedback. If the system is uncertain it will ask the student to self evaluate their reading.

To close the loop, when the feedback computed by the system is certain it is stored in the adaptive learning system as a correctness boolean. It is used to refine the student's profile and ensure the right level of exercise is given later on. Speech recordings are stored for continuous improvement purposes only. They are pseudonymized by design and anonymized at the end of each school year.

2.3. Teacher interface

Teachers can access a dashboard providing both an overview of the class achievements and individual reports on student progression and detailed mistakes. When a student plays the read-aloud exercise, the teacher can listen to the recording (up to 12 months) and visualize the automatic feedback on the target word(s) (in bold in the interface) as shown in Fig.2. Optionally, they can report an inaccurate feedback, which internally enables us to continuously improve the system's performance.

3. Conclusion

Lalilo is a web application that helps young children learn how to read at their own pace through differentiated instruction. It also helps teachers by providing data on their students' strengths and weaknesses. It allows students to practice read aloud in autonomy with immediate feedback provided by a reading mistakes detection system for French and English native speakers. Accurate speech recognition for children aged 5 to 8 is challenging, in particular for reading speech and in a classroom context. The Lalilo system implements an uncertainty estimation in order to avoid giving an inaccurate feedback to the students, which could be disengaging and damaging for their learning. A high uncertainty estimation leads to the student being asked to self-evaluate their reading, then receiving a neutral feedback. Teachers have access to their students readings and the automatic feedback provided by Lalilo, and can report inaccurate feedback for continuous improvement.

4. Acknowledgements

This project is funded by the company Lalilo, Subsidiary of Renaissance Learning, Inc.

5. References

[1] National Reading Panel (2000), "Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups." in *National Institute of Child Health and Human Development, NIH, DHHS*, 2000, pp. 00–4754.

[2] T. Shanahan, "The national reading panel report. practical advice for teachers." in *Learning Point Associates / North Central Regional Educational Laboratory (NCREL)*, 2005.

[3] J. Mostow and G. Aist, "Evaluating tutors that listen: An overview of Project LISTEN," in *Smart machines in education: The coming revolution in educational technology.* The MIT Press, 2001, pp. 169–234.

[4] D. Bolaños, R. Cole, W. Ward, E. Borts, and E. Svirsky, "FLORA: Fluent oral reading assessment of children's speech," *ACM Trans. Speech Lang. Process.*, vol. 7, no. 4, p. 16, 2011.

[5] E. Godde, G. Bailly, D. Escudero, M.-L. Bosse, and G. Estelle, "Evaluation of reading performance of primary school children: Objective measurements vs. subjective ratings," in *Proc. of the International Workshop on Child Computer Interaction (WOCCI)*, 2017, pp. 23–27.

[6] R. Mugitani and S. Hiroya, "Development of vocal tract and acoustic features in children," *The Journal of the Acoustical Society of Japan*, vol. 68, no. 5, pp. 234–240, 2012.

[7] A. Potamianos and S. Narayanan, "Spoken dialog systems for children," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 1, 1998, pp. 197–200 vol.1.

[8] G. Yeung and A. Alwan, "On the difficulties of automatic speech recognition for kindergarten-aged children," in *Proc. of the Annual Conference of the International Speech Communication Association (INTERSPEECH), Hyderabad*, 2018, pp. 1661–1665.

[9] L. Gelin, M. Daniel, J. Pinquier, and T. Pellegrini, "End-to-end acoustic modelling for phone recognition of young readers," *ArXiv preprint:2103.02899*, 2021.

[10] H. Jiang, B. Kim, M. Y. Guan, and M. Gupta, "To trust or not to trust a classifier," in *Proc. of the International Conference on Neural Information Processing Systems*, 2018, p. 5546–5557.