

My Turn To Read: An Interleaved E-book Reading Tool for Developing and Struggling Readers

Nitin Madnani Beata Beigman Klebanov
Anastassia Loukina Binod Gyawali John Sabatini
Patrick Lange Michael Flor

Educational Testing Service, Princeton, NJ, USA
{nmadnani, bbeigmanklebanov, aloukina, bgyawali,
jsabatini, plange, mflor}@ets.org

Abstract

Literacy is crucial for functioning in modern society. It underpins everything from educational attainment and employment opportunities to health outcomes. We describe My Turn To Read, an app that uses interleaved reading to help developing and struggling readers improve reading skills while reading for meaning and pleasure. We hypothesize that the longer-term impact of the app will be to help users become better, more confident readers with an increased stamina for extended reading. We describe the technology and present preliminary evidence in support of this hypothesis.

1 Introduction

According to the results of the 2017 National Assessment of Educational Progress (NAEP)¹, 32% of U.S. 4th graders read below the Basic level. Most such students lack foundational skills of oral reading fluency – accuracy, reading rate, and prosody. Furthermore, more than a million students *at* the Basic level are also relatively slow readers, have poor prosody, and make more errors than skilled readers (Sabatini et al., 2018).

The combination of low reading accuracy and slow reading rate likely take a toll on a young reader’s engagement and motivation to read. While there are many interesting fiction and non-fiction books available to young readers, a slow, laborious reading process can make the act of reading feel like work, not pleasure. The problem is perhaps most acute for children who do not have adults to read with them. Children who do not acquire text fluency in school are left to their own devices to try to bootstrap it without the feedback and motivation usually provided by a knowledgeable and supportive teacher or caretaker.

¹https://www.nationsreportcard.gov/reading_2017/nation/achievement?grade=4

My Turn To Read (MTTR) is an educational application designed to help such low-proficiency readers improve reading skills through sustained reading with technological support. To make the critical transition from word-by-word reading to fluency, readers need to be engaged in the flow and process of reading for meaning and pleasure, which cannot occur if getting through every page is a struggle. MTTR can be thought of as a virtual reading companion who narrates part of the story to help enhance engagement and alleviate frustration during reading, and ultimately to help improve confidence, fluency, and reading stamina.

In the next section, we describe the idea of interleaved or turn-based reading and its hypothesized benefits (§2). Next, we describe the MTTR app itself – its features, components, and any NLP & Speech technologies (§3). Next, we discuss the results of trialing MTTR with two summer camps (§4). We conclude with our future plans for MTTR for both additional features as well as additional NLP & Speech technologies (§5).

2 Oral Reading with Turn-Taking

Listening to and engaging in oral reading pervades daily life – parents reading aloud to children, children receiving reading instruction in schools, and adults choosing audio narration (for books & podcasts) as the reading medium that best fits busy schedules. Oral reading fluency is also an important indicator of reading skill (Fuchs et al., 2001).

The main idea behind interleaved book reading is to allow the user to take turns reading aloud from a long, challenging, high-interest book with a virtual partner, realized, in our case, through an audiobook narration. The text of the book is split into paragraphs which are then allocated to alternating narrator and user turns². During the

²The actual number of paragraphs in each turn may vary.

narrator turn, the user listens to the corresponding recording from the audiobook; during the user turns the user is prompted to read the text of the user turn aloud. The narrator and user turns do not overlap – the user continues reading from where the narrator left off, and vice versa.

We hypothesize that (a) the interest in the story and the quality of the narration increases enjoyment, and (b) the interleaving of effortful reading with the more relaxing experience of listening to a skilled narrator allows regular breaks for the user to rebuild stamina to continue reading. The combined effect of (a) and (b) is to make the process sufficiently easygoing and engaging for the user to continue reading the whole book with the app, thus gaining reading practice and boosting their skill, confidence, and enthusiasm as readers.

3 My Turn To Read App

In this section, we describe the current version of the MTTR application. The application is designed to be cross-platform – it works on the web as well as on the iOS and Android mobile platforms. It was particularly important to have mobile versions of the application since (a) it provides more flexibility to the users (i.e., kids can read on a computer or school tablet during school hours and continue reading on a different device at home) and (b) in our preliminary interviews with adult literacy learners – another target demographic of the app – a majority said that they used mobile phones as their *only* computing device.

Mobile versions of MTTR are built using Apache Cordova³ – a cross-platform toolkit – with platform-specific modifications where necessary. The reading and listening components in all versions are built on top of Radium⁴, a robust, standards-compliant, and open-source e-reader. Figure 1 shows a screenshot of the iOS version of MTTR.

As users read with MTTR, it logs information about their interactions. The audio from user turns is recorded and stored. The app also logs rich process data which allow reconstructing the timeline of a user’s interaction with the app, such as timestamps for the beginning and end of each user and narrator turn and the answers given to reading comprehension questions (see §3.2). Other than the turn audio, no other personally identifying in-

³<https://cordova.apache.org>

⁴<https://readium.org>

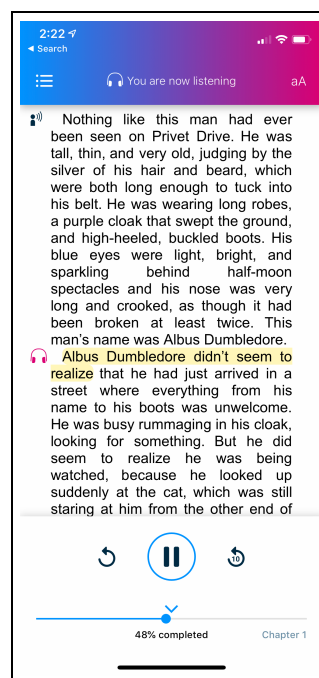


Figure 1: A screenshot of the iOS version of My Turn To Read (MTTR). Start of narrator turn is marked with the headphones & start of user turn is marked with speaker icon. Currently the narrator is reading (see status bar on top); the user can follow the narrator using the yellow highlight and also pause, replay, and rewind.

formation is collected and stored by the app. A separate, secure authentication server stores user-provided email used for registration. All collected data is stored in a secure database with strict access controls and no public access. The user is explicitly notified when the recording is about to start and via a status bar while it is in process.

Next, we describe the salient MTTR features along with the underlying NLP & Speech technologies, where appropriate. A video illustrating most of the user-facing features in action is currently available at <https://www.youtube.com/watch?v=Efs11ZMWFkE>.

3.1 Read Aloud eBooks

In order to use a book with MTTR, we need to combine the eBook and the audiobook versions of the book into a new format such that every paragraph is assigned a unique ID and the text is synchronized with the audio in the audiobook. These are necessary to (a) transition between listening and reading and (b) highlight text fragments in the eBook corresponding to the audio being played during narrator turns (as shown in Figure 1). The default highlighting is at the sentence level

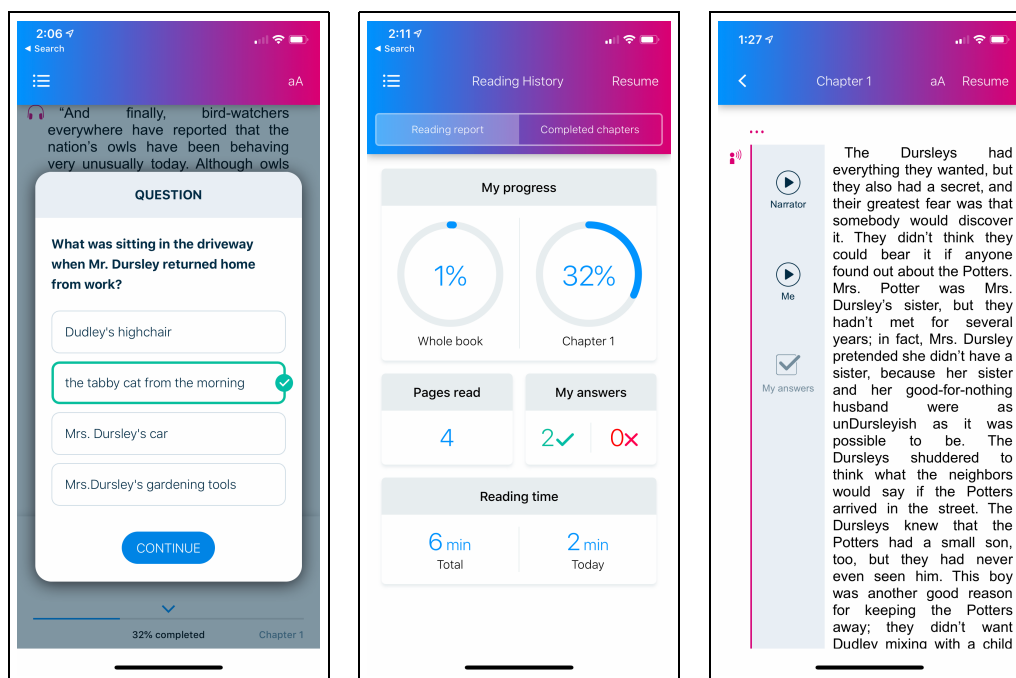


Figure 2: Screenshots illustrating reading comprehension questions and “Reading History” from MTTR: the first screen shows an example question, the second shows a report of the reading activity, and the third shows how readers can interact with already completed turns.

but we manually split long sentences into shorter spans based on syntax & narrator pauses and also make other adjustments to align with sometimes idiosyncratic narrator prosody. The purpose of the highlighting is to make it easier for a struggling reader to follow along during narrator turns, without the highlight moving so often as to be distracting (highlighting each word) or highlighting such large chunks of text as to defeat the purpose of closely following the narrator (highlighting complete sentences, no matter how long).

We use the EPUB format⁵ to create what we call a “Read Aloud eBook” used by MTTR. To link the text in the book to the synchronized audio, we use SMIL (Synchronized Multimedia Integration Language), as defined in the EPUB Media Overlays specification. The complete process for generating a Read Aloud eBook is as follows:

1. We use lxml⁶ to extract the plain text from the original eBook EPUB. We then break up paragraphs into sentences and create a mapping between sentence identifiers and token indices where sentences start and end.
2. We use forced alignment to align words in the normalized text of each chapter to the audio-

book MP3 file for this chapter. The alignment is done using the Kaldi ASR toolkit (Povey et al., 2011) and the LibriSpeech acoustic models (Panayotov et al., 2015). The resulting word-level alignment is used to compute the beginning and end timestamps for each sentence. We use Sequitur G2P (Bisani and Ney, 2008) to phonetically transcribe out-of-vocabulary words. The transcriptions are checked manually and added to the lexicon used for forced alignment.

3. We use ebooklib⁷ to generate a new EPUB file with sentences linked to time segments in the relevant MP3 file using SMIL.
4. We perform the splitting and other manual adjustments in the generated eBook to create subsentential highlighting spans as necessary. We then map any new spans back to the word-level alignment and regenerate the Read Aloud eBook with these spans as the highlighting units, linked via SMIL to audio timestamps. Subsentential spans can also be generated automatically (Parlikar and Black, 2012); we plan to use the manual splits to help improve automated splitting.

⁵<http://www.idpf.org/epub>

⁶<https://github.com/lxml/lxml>

⁷<https://github.com/aerkalov/ebooklib>

3.2 Reading Comprehension Questions

To check that users are paying attention to the story and to remind them of important story elements, we created approximately one reading comprehension question (RCQ) for every 100 words of running text. These are surface-level questions focused on the plot, on relationships between characters, on important descriptive details; the answers are usually stated in the text. Users are asked two questions after every other one of their turns. All questions are multiple choice with 2-4 options. Figure 2 shows an example.⁸

We also experimented with automated generation of RCQs using the semantic-role based system described in (Flor and Riordan, 2018). This system generated 1,350 questions for a 228-sentence excerpt from chapter 2 of *Harry Potter and the Sorcerer's Stone*. After removing all the questions that required resolution of pronominal or temporal anaphora to be sufficiently clear, as well as questions that contained incorrect information or were grammatically ill-formed, we were left with 280 questions for a closer examination. These questions were reviewed by an expert who has previously written RCQs used in the app. Of these, 75 (27%) were deemed usable as is or with a small fix (Q: "Why did Dudley have a tantrum?" A: "because his knickerbocker glory didn't have enough ice cream on top" illustrates Dudley's character; Q: "What did Uncle Vernon shout about once a week?" A: "that Harry needed a haircut" points at something unusual about Harry). Out of 280 questions, 150 (53%) were deemed unacceptable because they asked about a marginal detail (Q: "Who started looking for socks?" A: "Harry"). The remaining 20% of the questions had various problems such as insufficient specificity (Q: "Was Harry punished?" A: "no" requires more precise description of what he was or was not punished for in the particular instance in question), easily answerable based on general knowledge without reading the book (Q: "Who is slithering to the floor?" A: "the great snake"), awkward phrasing (Q: "What did Harry see?" A: "a huge Dudley tantrum coming on"), and too long to be readable (Q: "Had Dudley's gang been chasing him as usual when, as much to Harry's surprise as anyone else's, there he was sitting on the chimney?", A: "yes"). These findings suggest that above and

⁸Figures 1 and 3 are used by permission of the copyright owner Educational Testing Service.

beyond the known challenges of correctness of information and of form and non-anaphoricity, the biggest issue when generating questions based on a 100-word excerpt from a long story is choosing what to ask about. For MTTR, we want questions to also serve as reminders about important plot elements, characterizations, etc., and not just pick up on any minutiae.

3.3 Reading History

MTTR provides a section called "Reading History" containing two sub-sections. "Reading Report" allows users to keep track of how much they have read with MTTR (number of minutes that day and overall), what percentage of the current chapter (and the book) they have completed, and how many RCQs they have answered correctly. "Completed Chapters" allows users to revisit the turns completed so far: they can listen again to the narrator read its own turns and also listen to their own recordings of their turns. In fact, it also allows them to listen to the narrator read *their* turns since the audiobook contains narration for all paragraphs. Listening to themselves and then the narrator allows users to locate areas for improvement. This section also allows users to examine their answers to the questions that have been asked based on a given turn. Figure 2 shows the "Reading History" section from the app.

MTTR contains other useful features not described here in detail due to space limitations. For example, it allows users to adjust turn sizes – a really struggling reader might rely more on the narrator early on but gain the confidence to read aloud more as the book progresses. MTTR also allows readers to *re-record* their turns via "Reading History" if they catch some errors in their reading or get inspired by listening to the narrator.

4 Extrinsic Evaluation

In order for *any* reading app to have an impact on readers' skills – something that develops slowly and gradually – it is necessary for them to actually use the app consistently over a substantial period of time, preferably willingly.

We trialed MTTR with two summer camp programs in the greater NYC area in June–August 2018. One program ran for 6 weeks and included a reading session with the app for 20-50 minutes four days a week, with fewer days in the first week of the camp. The second program ran for a total

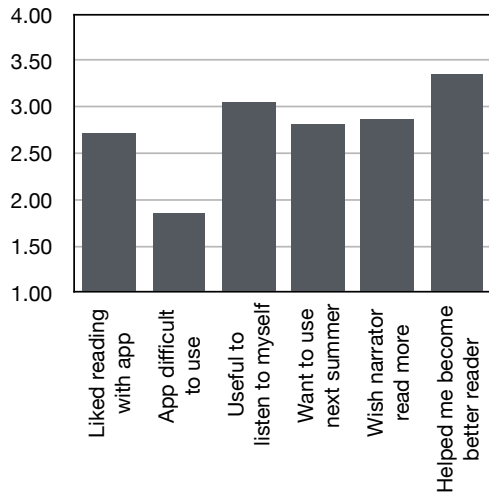


Figure 3: [Top] Children reading with MTTR in summer camps. [Bottom] Average ratings for MTTR survey questions from responses provided by 25 child participants in summer camp reading sessions. Each question was rated on a 4-point scale.

of 8 weeks (different children were enrolled for a different number of weeks) with a variable reading schedule depending on other camp activities; each reading session included about half an hour of reading and half an hour of related games and activities. All children read *Harry Potter and the Sorcerer’s Stone* by J.K. Rowling, with narration by Jim Dale. Children used MTTR on tablets connected to consumer-grade headsets with built-in microphones in a fairly laid-back, informal atmosphere; see Figure 3. A total of 36 children aged 8–11 participated in the two trials.

In both camps, children had the option to stop using the app entirely and engage in another camp activity. Of course, they could also hold the device but not actually *use* the app, or go through the motions of tapping on buttons but not actually do any listening or reading. We found that not only did children use the app when an opportunity was provided (based on the camp program), they also largely engaged with the app productively. In total, we logged more than 61 hours of listening (2,978 narrator turns). Our initial analy-

sis of user turns showed that 1,580 of them were of reasonable duration to make complete bona-fide reading of the turn possible (see (Beigman Klebanov et al., 2019) for details on estimating reasonable turn durations); based on transcriptions of these turns, they in fact contained 111 read words per turn on average. Finally, we also logged 9.5 hours spent answering 2,104 comprehension questions with 65% questions answered correctly. We also asked the children to fill out a survey at the end about their experience with MTTR. Figure 3 shows the the results from the 25 children who completed the surveys.

The fact that an overwhelming majority of the children who started reading with MTTR continued to use it for the duration of their camp enrollment *and* also continued to read aloud is a promising result. Furthermore, the positive responses to survey questions – particularly the one that asked if they believed that MTTR helped them become better readers – also suggest that MTTR has the potential to support extended reading and thus have the hypothesized positive impact.

5 Discussion & Future Work

My Turn To Read is currently in beta and we plan to release freely-available web⁹ and mobile (iOS & Android) versions in August of 2019 with the public-domain book *The Adventures of Pinocchio*. We plan to add more books in subsequent releases.

While the functionality implemented in MTTR has already yielded promising results, several avenues of future work are planned or underway.

We are already working on using automated speech recognition to track readers’ progress and provide useful automated feedback when appropriate (Loukina et al., 2017). Our plan is to first investigate a server-based speech processing system which will receive the readers’ speech over a (secure and encrypted) internet connection¹⁰. Based on our observations of the offline-vs-online usage and the latency profiles, we may decide that on-device speech processing is a better alternative.

We are working with users and teachers on determining what specific type of oral-reading-based feedback would be most useful (Kannan et al., 2019). Although automated processing of children’s speech holds promise for estimating read-

⁹<https://myturntoread.org>.

¹⁰MTTR stores recordings on-device until an internet connection is available.

ing skill, especially if we aggregate measurements from multiple user turns (Loukina et al., 2018; Wang et al., 2019), feedback for individual user turns is likely to be difficult due to substantial behavioral and technical noise in recordings, e.g., background noise, equipment malfunction, cross-speaker interference, skipped turns, mumbling, etc. (Loukina et al., 2018, 2019). Furthermore, we want to ensure that the feedback does not discourage already struggling readers (e.g., providing fluency scores may not be the right approach).

We plan to continue our work on automated question generation which will help shorten the turn-around time for adding new books.

Finally, we are exploring a use case for MTTR in classrooms in an ongoing trial with grade 3–5 students in an NJ elementary school. Although the results haven't been analyzed quantitatively, preliminary anecdotal evidence shows very positive reactions from both teachers and students.

Our goal is to help students thrive as fluent, confident, and enthusiastic readers; our hope is to be able to demonstrate quantitatively that MTTR can be instrumental in achieving this goal and, eventually, reduce the persistently high proportion of struggling readers in U.S. schools and elsewhere.

6 Acknowledgements

We thank the Astea Solutions team for the app design and development work; K. Dreier, V. Liceralde, J. Bruno, C. Appel, and I. Blood for creating the comprehension questions and K. Dreier also for her help with evaluating the automatically generated questions; J. Lentini for help with the summer camp data collection; Y. Qian and A. Misra for help with ASR. We also thank the site administrators and instructors in the two summer camps for implementing the summer reading program with MyTurnToRead.

References

Beata Beigman Klebanov, Anastassia Loukina, Nitin Madnani, John Sabatini, and Jennifer Lentini. 2019. Would you? Could you? On a tablet? Analytics of Children's eBook reading. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*.

Maximilian Bisani and Hermann Ney. 2008. Joint-Sequence Models for Grapheme-to-Phoneme Conversion. *Speech Communication*, 50(5):434 – 451.

Michael Flor and Brian Riordan. 2018. A Semantic Role-based Approach to Open-Domain Automatic Question Generation. In *Proceedings of the BEA Workshop*, pages 254–263.

Lynn S. Fuchs, Douglas Fuchs, Michelle K. Hosp, and Joseph R. Jenkins. 2001. Oral Reading Fluency as an Indicator of Reading Competence: A Theoretical, Empirical, and Historical Analysis. *Scientific Studies of Reading*, 5(3):239–256.

Priya Kannan, Beata Beigman Klebanov, Shiyi Shao, Colleen Appel, and Rodolfo Long. 2019. Evaluating Teachers' Needs for On-going Feedback from a Technology-based Book Reading Intervention. Presented at the Annual Meeting of the National Council on Measurement in Education.

Anastassia Loukina, Beata Beigman Klebanov, Patrick Lange, Binod Gyawali, and Yao Qian. 2017. Developing Speech Processing Technologies for Shared Book Reading with a Computer. In *Proceedings of the 6th International Workshop on Child Computer Interaction (WOCCI)*, pages 46–51.

Anastassia Loukina, Patrick Lange, Yao Qian, Beata Beigman Klebanov, Nitin Madnani, Abhinav Misra, and Klaus Zechner. 2019. The Impact of Ambient Noise on Measurement of Oral Reading Performance. Presented at the Annual Meeting of the National Council on Measurement in Education.

Anastassia Loukina, Beata Beigman Klebanov, Nitin Madnani, Abhinav Misra, Georgi Angelov, and Ognjen Todoc. 2018. Evaluating On-device ASR on Field Recordings from an Interactive Reading Companion. In *Proceedings of IEEE-SLT*.

Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. 2015. Librispeech: An ASR Corpus Based on Public Domain Audio Books. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.

Alok Parlikar and Alan W. Black. 2012. Modeling Pause-duration for Style-specific Speech Synthesis. In *Proceedings of Interspeech*.

Daniel Povey, Arnab Ghoshal, Gilles Boulianne, Lukas Burget, Ondrej Glembek, ..., and Karel Vesely. 2011. The Kaldi Speech Recognition Toolkit. In *Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*.

John Sabatini, Zuowei Wang, and Tenaha O'Reilly. 2018. Relating Reading Comprehension to Oral Reading Performance in the NAEP Fourth-Grade Special Study of Oral Reading. *Reading Research Quarterly*, 54(2):253–271.

Zuowei Wang, John Sabatini, and Tenaha O'Reilly. 2019. Harry Potter Knows How Well You Read: Estimating Children's Reading Ability from Oral Novel Reading. Presented at the Annual Meeting of the National Council on Measurement in Education.