

Motivation

In rural India, millions of children complete primary school every year without achieving even basic reading standards.

[ASER (2012)]

Reading research provides empirical evidence that assisted oral reading, with listening to a fluent reader, is very effective in improving a student's reading skills. Repeated readings of a passage have been shown to lead to improvements both in word decoding and in comprehension.

[Dowhower, 1994; Rasinski, 2003]

Digital technology that facilitates reading practice coupled with automated feedback is a powerful solution to the problems of teacher scarcity and widely differing learning paces. The objective feedback can also help educational authorities to track student progress.

Technology for Reading

Provide an engaging and immersive environment that encourages a child to attempt new material of pre-selected difficulty.

Facilitate assisted reading with a guiding voice and text captions where words are highlighted in sequence.

Encourage practice and self-assessment by enabling audio recording. The recordings can be available for teacher review.

Provide automated feedback based on oral reading rubrics.

Assessing reading skills

Two independent cognitive tasks occur in good reading: word decoding and comprehension.

Proficient readers organize the text into meaningful phrases and read with appropriate prosody and pace while pronouncing words correctly. Comprehension has been shown to be predictable by the prosody of the student's reading.

The scoring rubric therefore must incorporate (i) word decoding accuracy (word is missed/correct/incorrect), (ii) phrasing, (iii) volume and intonation, (iv) reading pace (medium/slow/fast).

In repeated readings, children typically start out by reading words in list-style with pauses and disfluencies. Gradually, they group words into phrases and apply correct intonation. Finally, they become accurate and expressive readers.



School in Dahanu tribal belt



Students engage with the Reading App



Pilot Testing

(March 2016 onwards)

Gram Mangal School in Dahanu (in Maharashtra's tribal belt)

Medium of Instruction: Marathi; English taught as Second Language

Oral Reading with Tab is a weekly activity with 30 min session per student

Number of children in Test Group: 80 across Classes 6, 7, 8, 9; age group: 10 - 14

Children have very limited exposure to Spoken English and no opportunities to speak

Automatic Assessment of Reading with Speech Recognition Technology

Preeti RAO, Prakhar SWARUP, Ankita PASAD, Hitesh TULSIANI and Gargi Ghosh DAS

Dept. of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, India



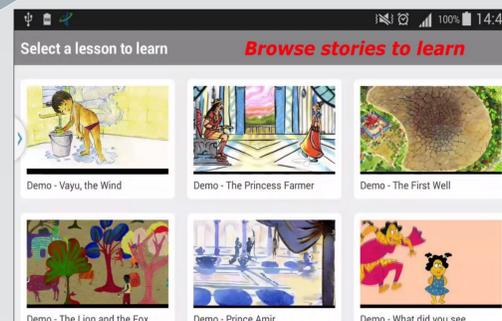
Everytime I finish my hot hot bath
My wet body feels so cool cool cool
Who makes that happen Vayu the wind
The milk in my cup too hot too hot
But soon it is ready for me to gulp
Who makes that happen Vayu the wind
The window curtains flutter and gently brush my face
Who makes it happen Vayu the wind
A bolt of lightning far away Black clouds moving my way
Who makes that happen Vayu the wind

Review screen with colour-coded feedback

Reading Tutor App [Sensibol] with Stories by BookBox on Android Tablet (7" screen) with mic-headset



Reading interface with text highlighting



Story catalog display

Automatic Speech Recognition (ASR)

Technology that converts acoustic speech signals to text using language and other constraints. ASR has been used previously in objective assessment of language skills.

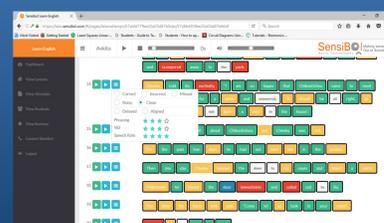
In the context of reading, ASR can ideally identify word-level errors. Speech-silence discrimination can help estimate speech rate. Pitch tracking can help predict phrasing and intonation deficiencies.

ASR is based on machine learning with substantial amounts of training speech data. Children's speech with its inherent variability, diversity of speaking accents and presence of background noise poses difficult challenges.

ASR training database creation

A task-specific labeled database is required to train acoustic and language models required by the ASR system.

Audio recordings obtained from the field-deployed tablets are uploaded to servers with their metadata. A text transcript is manually created for each child-story facilitated by a user-interface and backend programs that automatically segment the audio. Prosody ratings and noise labels are also obtained.



Semi-automatic ratings interface

Acknowledgements

Sensibol Audio Technologies for help with the customization of the Reading Tutor App for this study.

Prof. Alka Hingorani and her team at IDC, I.I.T. Bombay for initiating the LETS project and organizing the field study.

