

A hands-on demonstration of Project LISTEN's Reading Tutor and its embedded experiments

Jack Mostow, Greg Aist, Juliet Bey, Paul Burkhead, Andrew Cuneo, Susan Rossbach, Brian Tobin, Joe Valeri, and Sara Wilson

Project LISTEN, Carnegie Mellon University (www.cs.cmu.edu/~listen)
RI-NSH 4213, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA

Mostow@cs.cmu.edu

Abstract

Project LISTEN's Reading Tutor helps children learn to read. It uses speech recognition to listen to them read aloud, and responds with spoken and graphical feedback. The demonstration lets attendees try out this interaction themselves. Besides the spoken tutorial dialog, features shown include an automated tutorial for new users, interactive activities that combine assisted reading with other types of steps, and automated field studies to evaluate the efficacy of alternative tutorial interventions by embedding experiments within the Reading Tutor.

1 Introduction

Project LISTEN's Reading Tutor listens to children read aloud, and helps them learn to read (Mostow & Aist, 1999; Mostow & Aist, 2001). Compared to statistically matched classmates who spent the same time in regular classroom activities, children who used the Reading Tutor improved significantly more in comprehension in a 1998 study (Mostow et al, 2001b), and significantly more in vocabulary in a 1999-2000 study (Mostow et al, 2001a; Aist et al, 2001).

At least that is our goal. In point of fact, although guided oral reading has been shown to have educational value when done by properly trained human tutors (NRP, 2000), much human tutoring is not effective (Snow, 1998). And

although listening to children read has obvious face validity and some demonstrated success, realizing its full potential requires much more work. It is necessary, but not sufficient, to compare how well the Reading Tutor helps different children build different reading skills. We need to identify which elements of its rich, multimodal tutorial dialog actually help with which skills, which children, and which words.

In service of this long-range goal, we have embarked on a series of progressively more automated experiments, starting in 1996 with observational studies of students using the Reading Tutor. Along the way we have encountered several bottleneck issues:

- Experimenting, not just analyzing dialog
- Defining and computing outcome variables
- Viewing data so common sense can apply
- Harvesting, transporting, and pooling data
- Creating and communicating data analyses

We have refined our approaches to these issues, until we are now automatically collecting and analyzing field data from daily Reading Tutor use by hundreds of children at three schools.

One key idea is to embed experiments in the Reading Tutor (Aist & Mostow, 2000). Each experiment performs randomized trials to try alternative tutorial actions in a given situation. By aggregating over many trials by different children using multiple Reading Tutors, we can try to learn which actions work better, and when. Because all the trials are embedded in the same Reading Tutor, they control for Hawthorne effects that affect comparisons of the Reading Tutor to other treatments.

Automated field studies in 2000-2001 evaluated:

- Student ratings of thousands of stories read
- Transfer effects from oral to written spelling
- Alternative ways to preview new words
- Alternative ways to explain new vocabulary

2 Demo Outline

The demo traces the path from Tutor to results:

1. New reader enrolls him or herself.
2. New reader gets automated interactive tutorial about the Reading Tutor.
3. Reader picks and reads a story aloud with the Reading Tutor's help. Other activities include being read to, writing, narrating, and answering multiple-choice questions.
4. The Reading Tutor inserts preview and review activities before and after stories. Some of these activities implement embedded experiments. One such experiment compares ways to explain new words. Before a story it gives a definition for one new word, a synonym for another, a test on a third, and no help for a fourth. After the story it tests all four words.
5. Data from these embedded experiments are automatically logged and sent back every night from 21 Reading Tutors in three schools. Most data consists of log entries for trials, but one experiment also sends back recorded speech to transcribe by hand.
6. Data is parsed, aggregated, and analyzed in SPSS, with updated results viewable the next day via web browser.

The demo should include analyses based on hundreds of trials as recent as the day before.

Acknowledgements

This work was supported in part by the National Science Foundation under Grant Nos. REC-9720348 and REC-9979894, and by Gregory Aist's National Science Foundation Graduate Fellowship and Harvey Fellowship. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation or the official policies, either expressed or implied, of the sponsors or of the United States Government. We also thank the students and educators at Fort Pitt, Edgeworth, and Osborne Elementary Schools, and the rest of the Project LISTEN team.

References (see also www.cs.cmu.edu/~listen)

- Aist, G. and Mostow, J. (2000) Using Automated Within-Subject Invisible Experiments to Test the Effectiveness of Automated Vocabulary Assistance. In Joseph Beck (Ed.), *Proceedings of ITS'2000 Workshop on Applying Machine Learning to ITS Design/Construction*, pp. 4-8. Fifth International Conference on Intelligent Tutoring Systems, Montreal, Canada.
- Aist, G. S., Mostow, J., Tobin, B., Burkhead, P., Corbett, A., Cuneo, A., Junker, B., and Sklar, M. B. (2001) Computer-assisted oral reading helps third graders learn vocabulary better than a classroom control – about as well as one-on-one human-assisted oral reading. *Proceedings of the Tenth Artificial Intelligence in Education (AI-ED) Conference*, San Antonio, Texas, May 2001.
- Mostow, J. and Aist, G. (1999) Giving help and praise in a reading tutor with imperfect listening – because automated speech recognition means never being able to say you're certain. *CALICO Journal* 16:3, 407-424. Special issue (M. Holland, Ed.), *Tutors that Listen: Speech recognition for Language Learning*.
- Mostow, J. and Aist, G. (2001) Evaluating Tutors that Listen: An Overview of Project LISTEN. In K. Forbus and P. Feltovich (Eds.) *Smart Machines in Education: The coming revolution in educational technology*. MIT/AAAI Press. In press.
- Mostow, J., Aist, G. S., Burkhead, P., Corbett, A., Cuneo, A., Eitelman, S., Huang, C., Junker, B., Platz, C., Sklar, M. B., and Tobin, B. (2001a) A controlled evaluation of computer- versus human-assisted oral reading. In *Proceedings of the Tenth Artificial Intelligence in Education (AI-ED) Conference*, San Antonio, Texas, May 2001.
- Mostow J., G. S. Aist, C. Huang, B. Junker, R. Kennedy, H. Lan, D. Latimer, R. O'Connor, R. Tassone, B. Tobin, and A. Wierman. (2001b) 4-Month Evaluation of a Learner-controlled Reading Tutor that Listens. In Philippe DeCloque and Melissa Holland (Editors), *Speech Technology for Language Learning*. The Netherlands: Swets & Zeitlinger Publishers. In press.
- National Reading Panel. (2000) *Teaching Children to Read*. www.nichd.nih.gov/publications/nrppubskey.cfm
- Snow, C. E., Burns, M. S., and Griffin, P. (Eds.). (1998) *Preventing Reading Difficulties in Young Children*. National Research Council. Washington, DC: National Academy Press. At www.nap.edu/readingroom/books/prdyc.