Authoring New Material in a Reading Tutor that Listens

Jack Mostow and Gregory Aist

Project LISTEN

4910 Forbes Ave., LTI

Carnegie Mellon University

Pittsburgh, Pennsylvania 15213

mostow@cs.cmu.edu

Abstract

Project LISTEN's Reading Tutor helps children learn to read by providing assisted practice in reading connected text. A key goal is to provide assistance for reading any English text entered by students or adults. This live demonstration shows how the Reading Tutor helps users enter and narrate stories, and then helps children read them.

Areas: intelligent interfaces, computer-aided instruction, dialog, speech recognition

Why Authoring?

Project LISTEN's Reading Tutor listens to children read aloud. In an intelligent tutor for reading, why should students write? Writing as a part of intensive reading interventions such as Reading Recovery (Clay 1985) is believed to help students succeed at reading (Pinnell et al. 1994) and learning vocabulary (Gipe & Arnold 1978). Writing stories for other students to read can be motivational as well. In addition, students may more easily learn to read from stories written in familiar language styles (Serwer 1969), such as stories written by older schoolmates. Finally, allowing teachers to enter instructional material would allow the Reading Tutor to be more tightly integrated into the classroom.

Why narrate stories? The Reading Tutor eschews synthesized speech – used in Kurzweil's (1999) reading system – in favor of recorded human voices, which are much more expressive. Martin Luther King's stirring delivery of his "I Have a Dream" speech – an oft-chosen Reading Tutor selection – conveys this point dramatically. This preference induces a requirement to capture human narrations, especially in the author's own voice.

Writing a New Story

Adding a story starts with an Edit window (Figure 1) with initial text "My Story, by <FirstName> <LastInitial>." This box is a standard Windows editor, into which the user can type, or paste copied text.

The Edit window also gives young writers rudimentary spoken help. When the user types a letter, the Reading Tutor speaks it. When the user ends a word with a space or other punctuation, the Reading Tutor speaks the word. To accommodate variations in typing ability, the "talking typewriter" is speed sensitive. A letter or word is spoken only if the user hesitates before typing more. Thus only very slow typing (such as a child's) elicits letter names. Fluent typing with pauses between words speaks just the words. Rapid typing suppresses speech output.

Narrating the New Story

When the user leaves Edit after typing in or editing a story, the Reading Tutor enters narration mode (Figure 2). This mode differs from normal tutoring mode because its goal is to capture a fluent reading of the sentence, without substitutions, deletions, long hesitations, self-corrections, or other insertions. If the output of the speech recognizer matches the sentence perfectly, the Reading Tutor echoes the reading and goes on to display the next sentence. Otherwise, the Reading Tutor asks the user to read it again. This cycle repeats until the Reading Tutor accepts the sentence or the reader clicks *Go* (Figure 2) to proceed without narrating the sentence. At any time, the user can click *Back* to return to a previous sentence and re-record it.

Besides capturing sentence narrations, the Reading Tutor needs to capture individual words not previously recorded. The Reading Tutor uses the time alignment output by the speech recognizer to excerpt the segment of the recording corresponding to each unrecorded word. The time alignment is not always correct. Fortunately, the Reading Tutor already has high-quality recordings of hundreds of the most common words. Thus the remaining unrecorded words tend to be longer content words that get aligned more accurately than shorter words. To further reduce the effect of poor alignment, the Reading Tutor uses such captured word recordings only for the story where they are recorded, and not in other stories.

The *Edit Text* button displayed during narration mode lets the user return to the Edit window to modify the text. When the user leaves Edit, the Reading Tutor returns to narration mode, starting at the first unnarrated sentence.

Figure 1. Write (Edit).

Figure 2. Narrate.

Figure 3. Read.

The *Edit Media* button lets the user find a picture (using a standard file browser) to illustrate the current sentence.

When the user reaches the end of the story, the Reading Tutor leaves narration mode and returns to its menu of stories, which now include the narrated story.

Reading the New Story

The new story is now available for children to select and read. The Reading Tutor presents one sentence at a time, graying out earlier text (Figure 3). Listening with continuous, open-mike speech recognition, the Reading Tutor visibly shadows the word it expects to hear next, and tracks student performance, turning words green that it accepts as read correctly.

The Reading Tutor gives help on a word or sentence when the student clicks for help, gets stuck, makes a mistake, or is considered likely to misread a difficult word [Aist & Mostow, CALL97; Mostow & Aist, CALICO99]. The Reading Tutor may also backchannel after a brief silence, give praise for good or improved reading, go on to the next sentence when appropriate, or suggest what to click.

The Reading Tutor is designed to help the student read any input English text, making use of resources when they are available, and fallbacks when they are not. The narrated sentences and words are key resources.

For example, the Reading Tutor's most common intervention is to read a sentence aloud by playing its recorded narration. This intervention exploits the expressiveness of the human narration.

The time alignment captured in the narration process has several uses. The alignment lets the Reading Tutor highlight successive words as it reads a sentence. The alignment also allows the Reading Tutor to "recue" a word by rereading the words that lead up to it, and then underlining the word to prompt the student to reread it. The alignment lets the Reading Tutor extract in-context recordings of individual words from the sentence. This capability is especially useful when the Reading Tutor has no recording of the word in isolation. It also addresses the issue of homonyms (different words spelled the same) by providing the context-appropriate pronunciation to use.

If a sentence was not narrated, the Reading Tutor falls back on reading the sentence word by word. This intervention lacks expressiveness, but retains the quality of human speech. Finally, if a word is not recorded, the Reading Tutor uses a synthesizer to speak it. The Reading Tutor also uses a synthesizer to guess a pronunciation to listen for in the speech recognizer if a word is not in the pronunciation dictionary.

Word help may include saying the word, recuing the word, sounding or spelling it out, splitting it (visibly and audibly) into syllables, giving a rhyming hint, or (if available) displaying a picture or playing a sound effect.

Acknowledgements

References

(For Project LISTEN publication references, please see the list of publications at http://www.cs.cmu.edu/~listen.)

Clay, M. 1985. *The Early Detection of Reading Difficulties.* (Third ed.) Portsmouth, NH: Heinemann.

Gipe, J.P., and Arnold, R.D.. 1978. Teaching vocabulary through familiar associations and contexts. Journal of Reading Behavior 11(3): 281-285.

Kurzweil Educational Systems, Inc. 1999. Kurzweil 3000. http://www.kurzweiledu.com/kurzweil3000.html

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., and Seltzer, M. 1994. Comparing instructional models for the literacy education of high-risk first graders. Reading Research Quarterly 29(1), pp. 8-39.

Serwer, B. L. 1969. Linguistic support for a method of teaching beginning reading to black children. Reading Research Quarterly 4(4): 449-467, Summer 1969.