Reading Companion: A Interactive Web-Based Tutor for Increasing Literacy Skills

Keith Grueneberg[1], Amy Katriel[1], Jennifer Lai[1], and Jing Feng[2]

[1] IBM T.J. Watson Research, 19 Skyline Drive Hawthorne, NY 10532
[2] Fordham University, New York, NY 10023
{kgruen, amygk, jlai}@us.ibm.com, jfeng@fordham.edu

Abstract. This paper discusses learnings from the development of Reading Companion, a web-based system that uses speech recognition technology to help children and adults increase their literacy skills. An animated tutor character both 'speaks' and 'listens' to the reader, guiding them as they read e-books selected by their teacher from a virtual library. The system creates detailed performance reports for each student and because it is available on the internet, the learners can continue reading where they left off once they get home to share their progress with their family. We discuss implementation challenges that were overcome, as well as feedback from users and teachers. To our knowledge this is the first successful implementation of real-time interactive speech recognition using Flash on the internet. We believe this presents a valuable model of how speech can be used on the web as part of interactive applications.

Keywords: Literacy, web, speech recognition, animated tutor character.

1   Introduction

Approximately 90 million adults in the United States have limited literacy skills [4]. The key to increasing literacy skills is practice, and successful practice requires immediate feedback. The concept of using speech recognition technology as an aide to improving reading skills has been around for a decade; however none of the earlier solutions were available through secure access on the worldwide web. Previous applications [e.g. 1, 2, 5] required CD-ROM installation of the software. Watch Me! Read, [3] was the precursor to Reading Companion, using speech recognition with an interactive tutor to lead the student through the process. Like other solutions of its ilk, it required a CD-ROM installation. A lack of centralized results required teachers to go to each computer in the classroom to view student results, and for each student to use the same machine every time. Further, students could not continue their work at home and have those results appended to the results acquired at school. Reading Companion (RC) is a web-based literacy solution that addresses these problems. It has recently been made freely available through grants to not-for-profit organizations such as elementary schools and adult learning centers, operating in over 200 sites with almost 10,000 registered users.

2 Reading Companion

Reading Companion engages students with an animated character who provides help-ful instruction on how to interact with the e-book (e.g. "ok, now you read the words in red") and 'listens' to the reader using speech recognition technology, providing real-time feedback to emerging readers (e.g. "you sound great"). A screen shot of the ap-plication is shown in Figure 1. Automatic changes in scaffolding levels, where the tutor reads more for novice readers and less for more advanced readers, make the same books accessible for different levels of readers.
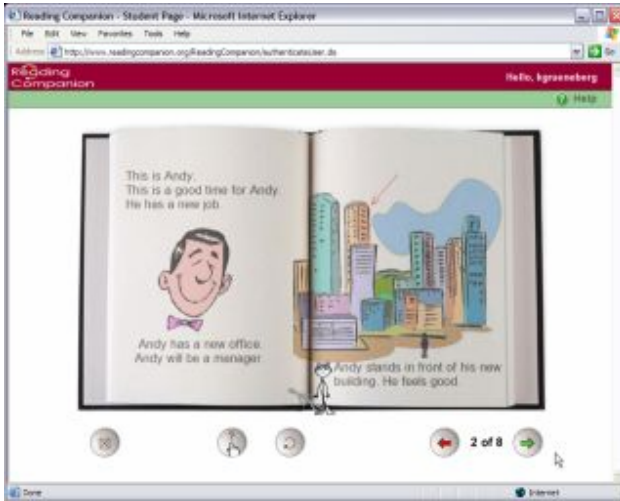

Fig. 1. An online e-book for adults

Speech recognition over the web can be accomplished either by deploying a speech recognizer on a server and sending the compressed speech over the network, or by embedding the speech capabilities in the browser and doing the decoding locally. Server-based speech recognition systems are optimized for telephony applications (i.e. automated phone answering systems), but do not work as well for web browser-based applications. Calls from a web client to a server-based recognizer are subject to latency problems due to the cost of transmitting audio over a network. An initial ver-sion of Reading Companion (RC) that used a server-based recognizer revealed laten-cies in schools that varied from 1 to 32 seconds. As a result, we had to re-architect the system to use an embedded speech recognizer rather than a server-based recognizer. RC uses a lightweight VoiceXML browser that is launched from Internet Explorer (or Firefox), which allows communication directly from Flash to the speech recognizer. This browser, runs on top of the IBM Embedded ViaVoice engine. To our knowledge this is the first time speech recognition has been embedded into a browser and used interactively on the web as part of a Flash application.

When a user accesses Reading Companion for the first time, they are prompted to download the speech software using a browser plug-in. The speech recognizer runs in the background as a student is reading. Because the application knows what words are

displayed on the page that is being presented, it can load a dynamic vocabulary, which creates very accurate speech recognition even for people with heavily accented English. Sample data collected in the field, showed a word error rate of 6.7% for the adult speech engine, and 4.6% for the child engine. Currently, when the companion 'speaks' the audio consists of pre-recorded human speech. This speech is recorded at the time the book is created. All speech recognition takes place on the client and the data for RC is stored on a centralized server. The DB2 database stores information on each student's performance, registration data, classroom settings, user session data (e.g. where the student left off in the book), as well as the book library.

The teacher, or a computer admin person, usually sets up RC on all the machines since first and second grade students are not sufficiently literate, in either computer or reading skills, to follow even the simple directions (e.g. click here to download). Each student had his/her own headset with microphone, and they log in by picking their name from a list of students. The reader (either adult or child) is presented with a library of books that the teacher has pre-selected for the class. Usually the teacher will also direct the student which book to pick since often these correspond to concepts or vocabulary that is currently being taught in class. Adult learners are presented with books geared toward things like finding employment, getting a driving license or gaining American citizenship. They receive feedback reading in English, obtain the necessary vocabulary to find a job, and increase their computer skills (e.g. mouse usage). Since RC is available on the internet, students can continue to access it from home. Reading Companion saves the user's session on the web server, so they can continue reading wherever they left off in school.

3   Feedback from Users

Similar to reading with a parent or teacher, the companion character provides support by modeling words (i.e. reading to the student), highlighting the text to be read, and giving both encouragement and immediate feedback. We use a panda character while the adults now interact with a stick-like figure (see Figure 2). Initial software trials used a character for adults shaped like a microphone who 'spoke' even though it did not have a mouth and it said things like "Click on my picture to start reading". It also 'listened' even thought it had no ears, and thus suffered from a credibility problem. We observed that the majority of our adult immigrant users would click on all pictures in the book in response to the instruction "click on my picture", without realizing that the microphone was talking about itself! The switch to a character which was more suggestive of human characteristics (such as ears and a mouth) alleviated all traces of this confusion.

Fig. 2. Adult, child and microphone characters

Tutoring applications need an effective way of modeling the student's progress, so that the system can adapt to the student's current skill level. The RC student model incorporates an extensive dictionary of words that are mapped to sound spellings (e.g. the silent "e" rule) and word features (e.g. consonant blends). The system stores results for each student in a database, and as the reader improves, less scaffolding is provided. The scaffolding levels are automatically adjusted based on the progress of the reader and we had to modify the highest level based on feedback from teachers who wanted additional room for growth for the most proficient readers.

Reading Companion provides detailed reporting for teachers, giving them visibility into the reader's areas of strength and weakness. The top of the report contains a summary of the overall accuracy of the student's reading, along with a visual sequential representation of all the words read by the student. The correct words are shown in green and the incorrect words are shown in red, thus allowing large patterns to be easily detected. As the teacher mouses over a red or green line, a popup shows the details of the word. Teachers really liked the visual display and the summary information since they often did not have the time to read the detailed findings for each student, but wanted to get an overview of the student's progress.

4   Conclusion

This paper describes the architecture, rationale and findings for a web-based literacy application called Reading Companion that is used by both adults and children. Feedback from users suggests that using speech recognition over the web to help learners increase reading skills is not only a viable solution, but one that is well accepted by the user populations. With RC we introduce a method that can be used by other interactive applications that wish to use speech technologies on the web.

There are many advantages to having a literacy tutor available over the web. The centralized web server makes it easy to administer and update the software, which is a huge advantage in terms of maintenance and adding new books to read. From the user's standpoint, the web provides anytime/anywhere access and a familiar browser interface. The ability to access Reading Companion from home allows the family members to stay involved with the reader's progress, and the learner can access the same content at home as in the classroom. Lastly, with RC we are able to impact emerging readers around the globe – all joining and building our web-based community of learners, teachers, and book-authors.

References

1. Mostow, J., Beck, J.E.: Project LISTEN's Reading Tutor: Interactive Event Description. In: Supplemental Proc. AIED 03, pp. 30–32 (2003)
2. Price, P., Luc, J.: Using speech and language technology to coach reading. In: Proc. HLT '01, Association for Computational Linguistics, pp. 1–2 (2001)
3. Williams, S.M., Nix, D., Fairweather, P.: Using Speech Recognition Technology to Enhance Literacy Instruction for Emerging Readers. In: Proc. ICLS 2000, pp. 115–120. Erlbaum, Mahwah (2000)
4. http://www.nifl.gov/nifl/facts/facts_overview.html
5. http://soliloquylearning.com/