The Automated Text Adaptation Tool

Jill Burstein, Jane Shore, John Sabatini, Yong-Won Lee & Matthew Ventura
Educational Testing Service
Rosedale Road MS 12R
Princeton, New Jersey 08541
{jburstein, jshore, jsabatini , ylee, mventura}@ets.org

1. Introduction

Text adaptation is a teacher practice used to help with reading comprehension and English language skills development for English language learners (ELLs) (Carlo, August, McLaughlin, Snow, Dressler, Lippman, Lively, & White, 2004; Echevarria, Vogt and Short, 2004; Yano, Long and Ross, 1994). The practice of text adaptation involves a teacher's modification of texts to make them more understandable, given a student's reading level. Teacher adaptations include text summaries, vocabulary support (e.g., providing synonyms), and translation. It is a time-consuming, but critical practice for K-12 teachers who teach ELLs, since reading-level appropriate texts are often hard to find. To this end, we have implemented the Automated Text Adaptation Tool v.1.0 (ATA v.1.0): an innovative, educational tool that automatically generates text adaptations similar to those teachers might create. We have also completed a teacher pilot study. Schwarm and Ostendorf (2005), and Heilman, Collins-Thompson, Callan, and Eskenazi (2006) describe related research addressing the development of NLP-based reading support tools.

During our interactive demonstration, conference participants can (a) login to the Internet-accessible tool, (b) import text files, and (c) experiment with adaptation features. We are currently interested in feedback from the computational linguistics community to inform tool development related to (a) feature enhancement, and (b) ideas for new NLP-based features. Until now, our primary source of feedback has been from teachers toward tool development from an educational perspective.

2. The Automated Text Adaptation Tool

NLP-based text adaptation capabilities in the tool are described in this section (also see Figure 1.) These adaptation features were selected for implementation since they resemble teacher-based adaptation methods.

2.1 English and Spanish Marginal Notes

Pedagogically, marginal notes are a kind of text summary. The *Rhext* automatic summarization tool (Marcu, 2000) is used to produce marginal notes in English. The amount of marginal notes generated can be increased or decreased based on students' needs. Using Language Weaver's[1] English-to-Spanish machine translation system, English marginal notes can be translated into Spanish.

2.2 Vocabulary Support

Synonyms for lower frequency (more difficult) words are output using a statistically-generated word similarity matrix (Lin, 1998). ATA v.1.0 generates *antonyms* for vocabulary in the text using WordNet®.[2] *Cognates* are words which have the same spelling and meaning in two languages (e.g., *animal* in English and Spanish). The tool generates these using an ETS English/Spanish cognate lexicon.

2.3 English and Spanish Text-to-Speech

The tool offers English and Spanish text-to-speech (TTS)[3]. English TTS may be useful for pronunciation support, while Spanish TTS provides access to the Spanish texts for Spanish-speaking ELLs who are not literate in Spanish.

---

[1] See http://www.languageweaver.com
[2] See http://wordnet.princeton.edu/
[3] See http://www.cstr.ed.ac.uk/projects/festival/ & http://cslu.cse.ogi.edu/tts/download/.
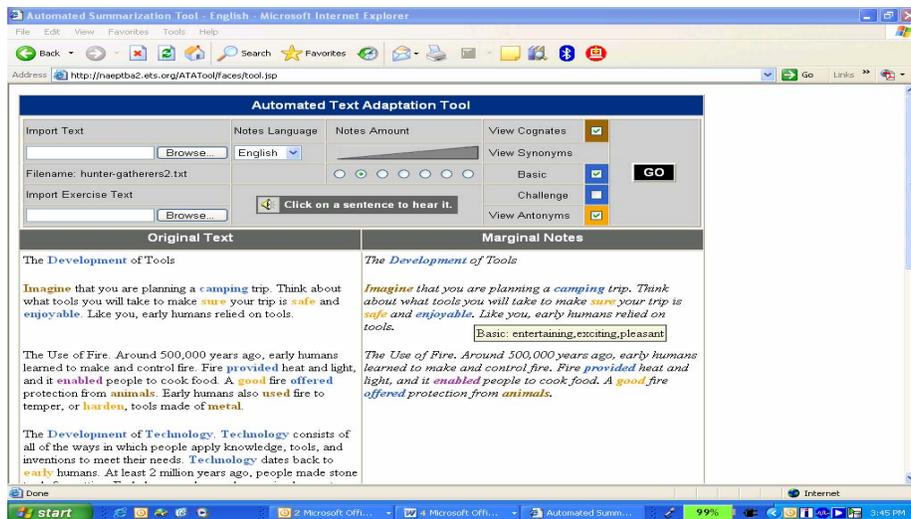
**Figure 1. Example Main Interface Screen showing *English Marginal Notes*
in the right column and *Synonyms* for *"enjoyable"* (*entertaining*, *enjoyable*, *pleasant*.)**

3. Pilot Study with Teachers

The survey feedback indicated that the 12 teachers were positive about the tool's potential. Overall, the vocabulary and English marginal notes were the most favorite features, while the text-to-speech was the least favorite. Teachers commented that they would like to see an editing capability added that would allow them to make changes to the automatically generated outputs (i.e., vocabulary support, and English and Spanish marginal notes.) Teachers viewed the tool either as *lesson planning support*, or as a *student tool for independent work*.

4. Future Research

ATA v.1.0 is a young application that uses NLP methods to create text adaptations. The teacher pilot evaluation suggested that it produces adaptations with potentially effective support for ELLs. It could also save teachers lesson planning time. We are currently implementing teacher-suggested modifications, and planning a larger, school-based pilot. The pilot will evaluate the tool's effectiveness in terms of measurable learning gains in reading comprehension and English language skills.

References

Carlo, M.S., August, D., McLaughlin, B., Snow, C.E., Dressler, C., Lippman, D., Lively, T. & White, C.

(2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream classrooms. Reading Research Quarterly, 39(2), 188-215.

Echevarria, J., Vogt, M., and Short, D. (2004). *Making Content Comprehensible for English Language Learners: the SIOP model.* New York: Pearson Education, Inc.

Heilman, M., Collins-Thompson, K., Callan, J., Eskenazi, M. (2006) Classroom Success of an Intelligent Tutoring System for Lexical Practice and Reading Comprehension. *In Proceedings of the Ninth International Conference on Spoken Language Processing.* Pittsburgh.

Lin, D. (1998). Automatic Retrieval and Clustering of Similar Words. *In Proceedings of the 35th Annual Meeting of the Association for Computational Linguistic*s, Montreal, 898-904.

Marcu, D. (2000) *The Theory and Practice of Discourse Parsing and Summarization.* The MIT Press, Cambridge, Massachusetts.

Schwarm, S. and Ostendorf, M. *Reading Level Assessment Using Support Vector Machines and Statistical Language Models.* In Proceedings of the Association for Computational Linguistics, Ann Arbor, MI, 523-530.

Yano, Y., Long, M. & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning, 44*, 189-219.