

# 2007 Student Research Symposium

*Language Technologies Institute*



All in the (word) family: Using learning  
Decomposition to Estimate Transfer Between Skills in a Reading Tutor that  
Listens

Xiaonan Zhang

## **Abstract:**

The work is carried out in the context of Project LISTEN's Reading Tutor, an intelligent tutor that helps students learn to read with Automatic Speech Recognition techniques. The overall goal of the Reading Tutor is to improve literacy by providing students with a pleasant, authentic experience in assisted reading.

As part of the project, I mainly work on discovering useful patterns in students' learning, by applying machine learning techniques to Reading Tutor's huge collection of data. It would enable us to build better student models and to improve tutor strategies.

The work focuses on one aspect of student modeling: estimating the transfer from learning one skill to similar skills. I'm interested in this topic because the amount of transfer reflects the student's mental representation of the skill, and also affect the benefits of practice. In this work specifically, I study transfer in the context of reading, and try to investigate whether the skill gained from exposure to a word transfers to similar words. Here "similar" is defined as "sharing the same word root".

I treat every word as a distinct skill, and distinguish between 4 non-overlapping types of practice opportunities for a word  $w$ :

D (Distributed same-word exposure): see word  $w$  for the first time that day.

M (Massed same-word exposure): see  $w$  again on the same day.

DS (Distributed similar-word exposure): see its root for the first time that day, in some word similar to  $w$ .

MS (Massed similar-word exposure): see its root again that day, in some word similar to  $w$ .

Finally I use learning decomposition to estimate the influence of each type of practice opportunities, from the student's word-reading performance and history of practice on that word. The influence of practice on similar words relative to the practice on same words gives a measurement of the transfer from practicing one word to learning similar words.

The data for evaluation was collected by the Reading Tutor during school year 2003-2004, including 860,517 observations from 346 students from Grade 1 to Grade 5.

First, I build a learning decomposition model for each student, and took the medians of all individual models as the overall estimates. The overall model shows that there is significant transfer from learning similar words; and distributed practice of similar words is even more effective than seeing the word itself again on the same day. In contrast, the impact of massed practice on similar words is negligible.

To investigate whether there is a positive relationship between initial reading skill and the amount of transfer, I divide all the students into low, medium and high proficiency, according to their pretest score. Then I calculate the median parameters for each subgroup. The result supports my initial hypothesis that transfer increases with the growth of students' reading skills. It also indicates that proficient readers benefit almost as much from practice on similar words as they do from the word itself.

In a similar way, I analyze the relationship between learning gains and the amount transfer, dividing the students into 3 bins based on their improvement over the school year. However, none of the differences between bins turn out to be significant.

These results validate our method as a powerful way to exploit the massive and detailed machine observations of students' activities (especially outputs from the Automatic Speech Recognition) to make quantitative inference about students' cognitive process.

Zhang, X., Mostow, J., & Beck, J. E. (2007, July 9-13). *Can a Computer Listen for Fluctuations in Reading Comprehension?* Proceedings of the 13th International Conference on Artificial Intelligence in Education, Los Angeles, CA.