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

Greg Aist, Jack Mostow, Brian Tobin,
Paul Burkhead, Albert Corbett, Andrew Cuneo, Brian Junker, and Mary Beth Sklar
Project LISTEN, Robotics Institute,
Carnegie Mellon University, Pittsburgh PA 15213 USA

Abstract. We describe results on helping children learn vocabulary during computer-assisted oral reading. This paper focuses on one aspect – vocabulary learning – of a larger study comparing computerized oral reading tutoring to classroom instruction and one-on-one human tutoring. 144 students in second and third grade were assigned to one of three conditions: (a) classroom instruction, (b) classroom instruction with one-on-one tutoring replacing part of the school day, and (c) computer instruction replacing part of the school day. For second graders, there were no significant differences between treatments in word comprehension gains. For third graders, however, the computer tutor showed an advantage over classroom instruction for gains in word comprehension ($p = 0.042$, effect size = 0.56) as measured by the Woodcock Reading Mastery Test. One-on-one human tutoring also showed an advantage over classroom instruction alone ($p = 0.039$, effect size = 0.72). Computer tutoring and one-on-one human tutoring were not significantly different in terms of word comprehension gains.

1. Introduction

This paper addresses an indispensable skill by using a unique method to teach a critical component: helping children learn to read by using computer-assisted oral reading to help children learn vocabulary.

Reading runs deeper than merely turning print into sound; reading makes meaning from print. Children must acquire a wide range of skills to ultimately comprehend text [1, 2]. Phonemic awareness allows children to distinguish and manipulate individual sounds in spoken words. Knowledge of print conventions enables children to work with text as placed on a page – for English, left-to-right, top-to-bottom. Mastery of the alphabetic principle reveals that individual sounds are written with letters or letter patterns. Decoding skills codify how to turn printed letters into sounds. Increased fluency leads to faster and more automatic reading. Background knowledge increases text understanding. Vocabulary knowledge is critical for comprehension. Drawing inferences from text and integrating information from multiple sources finally allow the reader to make meaning from print.

We focus in this paper on learning vocabulary. But what does it mean to know a word? A person's knowledge of a word may range from none at all to complete mastery. Aspects of word knowledge include:

- *Pronunciation. astronaut* is pronounced [æ s t ɹ o n ɔ t] in the International Phonetic Alphabet.
- *Spelling. astronaut* is spelled A S T R O N A U T.
- *Part of speech. astronaut* is a noun.
- *Morphology*. Inflectional morphology carries agreement. For example, the plural of *astronaut* is *astronauts*. Derivational morphology turns one word into another. For example, *astronaut* (noun) + *-ic* → *astronautic* (adjective).
- *Syntax*. For example, *astronaut* refers to a person, so the word *astronaut* takes he or she as a pronoun.
- *Lexical semantics*. The core meaning of *astronaut* is a space traveler.
- *Pragmatics*. For example, *cosmonaut* shares the core meaning of *astronaut*, but is used for Russian astronauts.

In this paper, we focus on learning the core meanings of words.

Assessing vocabulary knowledge is a difficult problem because we must sample along two dimensions: first, we must select a subset of words to test out of all the words in English – either focusing in on a particular kind of word, or representatively sampling from all words. Second, we must select a subset of aspects of word knowledge to test. For example, we could select 20 words at random from a dictionary, and then choose to test children's ability to define those words or use them in a sentence.

In this paper, we use a well-known published measure of vocabulary knowledge, the Word Comprehension subsection of the Woodcock Reading Mastery Test [3].

How can we help children learn new words? We consider two primary methods: direct instruction and learning through reading; and a hybrid: adding information to text.

1.1. Learning the meaning of a word through direct instruction

Intensive study of specific vocabulary words results in solid knowledge of the taught words, but at a high cost in time. For example, a 1983 study taught fourth graders 104 words over a five-month period, with 75 lessons of approximately 30 minutes each – on average about 21 minutes of instructional time per target word [4]. Exposures were during varied tasks: "matching words and definitions, associating a word with a context, creating contexts for words, and comparing and contrasting words to discover relationships" [4]. In the high-exposure group of words, students saw 26-40 exposures; even for the low-exposure words, students saw 10-18 exposures – a substantial amount of instructional time. Beck and McKeown [5] suggest that "the problem that effective instruction takes time can be alleviated by targeting instruction toward the most useful words" [5]. Which words are the most useful? They suggest second-tier vocabulary [6], that is, words that are "of high frequency in a mature vocabulary and of broad utility across domains of knowledge" [5]. Thus, direct instruction may play a role for certain critical words [7]. Nonetheless, a full-fledged instructional lesson is too time-consuming to use for every new word.

1.2. Learning the meaning of a word through reading

Children can learn words from written contexts [8, 9, 10], but the process is incremental. That is, the amount learned from each exposure may be small, but the net effect is still substantial [11]. Also, readers with better vocabularies learn more from context – because of broader and deeper prior knowledge of words – even though less of the information in the text is new to them than to readers with poorer vocabularies [12].

Reading offers hope for spurring vocabulary growth – if children can be guided to read material that does in fact contain unknown words. Carver [13] argues that "students must read books above their independent level in order to consistently run into unknown words, that is, about 1, 2, or 3 unknown words for each 100 words of text". Easier text simply does not contain enough new words to substantially improve children's vocabulary [13].

Is simple exposure to text sufficient for all readers to learn new words? Perhaps – or perhaps not. McKeown [9] studied how high- and low-ability students learn words from context. McKeown's [9] study examined 15 fifth-graders who, at the end of fourth grade, had scored between grade equivalent 3.3 and grade equivalent 4.1 on the Vocabulary section of the Stanford Achievement Test [14]. These low-reading fifth graders had trouble learning words from context partly because of incorrect inferences about the meaning of a word from context. One might expect that multiple sentence contexts would eliminate incorrect inferences – not the case. Both the low-reading fifth graders and the 15 higher-ability students in McKeown's 1985 study, who had scored above grade equivalent 4.8 on the Stanford Vocabulary subtest, had some trouble integrating multiple sentence contexts to derive meaning.

There has been some work aimed at teaching children how to learn words from context, but the major effect may be due to practice at learning new words from context and not due to teaching a specific strategy [15]. Kuhn and Stahl conclude that "Ultimately, increasing the amount of reading that children do seems to be the most reliable approach to improving their knowledge of word meanings, with or without additional training in learning words from context" [15]. As Schwanenflugel, Stahl, and McFalls [16] put it, "… the vast majority of a person's word growth can be accounted for by exposure to words in written and oral contexts, not through direct instruction of some sort, but individual encounters with a word in a natural context are not likely to yield much useful information about that word."

1.3. A hybrid: Adding information to text

Can the context in which a word appears be augmented in some way to make it more useful for learning the word? Typical dictionary definitions may not be written to suit the learner's needs; explanations written to convey the core sense of the word in plain language work better [6]. Presenting context-specific definitions in computer-mediated text has been shown to be helpful for vocabulary acquisition, at least for sixth graders [17]. Adding information to text is a hybrid of direct instruction and learning from reading text: first, start with a text to read; second, add brief, targeted instruction about words to the text. We adopted such a hybrid method – adding information to text – as described below in Section 1.3.

Now we have discussed what it means to know a word, and how to assess vocabulary knowledge. We now focus in on the specific area of our paper. In this paper we investigate learning words by reading connected text – including extra vocabulary assistance – during computer-assisted oral reading. We focus on encountering a word for the first time, and on learning the meaning of a word.

1.4. Learning vocabulary from assisted oral reading

In this section we describe the process of learning vocabulary during assisted oral reading. We describe an informal model: a conceptual framework useful for identifying opportunities to improve vocabulary learning. We started with Project LISTEN's Reading Tutor, a computer tutor that listens to children read aloud and helps them learn to read. The Reading Tutor displays a story one sentence at a time, and listens to the student read all or part of the sentence aloud. The Reading Tutor provides help in response to student mouse clicks and its analysis of the student's reading. Help available includes reading all or part of the sentence aloud, sounding out a word either sound-by-sound or syllable-by-syllable, and providing a rhyming hint. We can characterize how many words a student learns in a day of assisted oral reading as shown in

$$\frac{\text{New words learned}}{\text{Day}} = \frac{\text{Time reading}}{\text{Day}} \times \frac{\text{Stories read}}{\text{Time reading}} \times \frac{\text{New words seen}}{\text{Story read}} \times \frac{\text{New words learned}}{\text{New words seen}}$$

Equation 1. New words learned per day of assisted oral reading.

In other publications we have described two enhancements to the 1997-1998 Reading Tutor aimed at improving the number of new words learned per day. To increase the amount of new material that students read (and thus, increase new words seen per story read), we changed the Reading Tutor's story choice policy to have the computer and the student take turns picking stories (Take Turns). We showed that taking turns resulted in students reading more new material than they presumably would have on their own [18, 19]. In order to increase new words learned per new words seen, we enriched text with vocabulary assistance in the form of automatically generated factoids like "astronaut can be a kind of traveler. Is it here?" Aist [20] described how we showed that factoids helped, at least for third graders seeing rare words, and also for single-sense rare words tested one to two days later.

So, the changes we made improved the baseline 1997-98 Reading Tutor – at least in terms of the subgoals shown in Equation 1. But how did the new and improved Reading Tutor with Take Turns and factoids compare to other methods of helping children learn to read? Specifically, how did the 1999-2000 Reading Tutor compare to other reading instruction, on measures of vocabulary learning? In this chapter we present relevant parts of a larger 1999-2000 study that compared the Reading Tutor with classroom instruction and also with one-on-one human tutoring.

Project LISTEN conducted a year-long study in 1999-2000, comparing three treatments: the Reading Tutor, human tutoring in the form of assisted reading and writing, and a control condition with equal-time classroom instruction (not necessarily in reading). The primary purpose of the year-long study was to "prove and improve" the Reading Tutor: to compare the Reading Tutor to conventional methods of reading instruction, and to identify areas for improvement. This paper focuses on vocabulary learning; therefore we report here only the parts of the story relevant to vocabulary learning.

We now summarize the experimental design.

2. Experimental design

A total of 144 students in grades 2-3 at an urban elementary school near Pittsburgh, Pennsylvania began the 1999-2000 study, with 131 completing the entire year including post-testing. Each student received one of (a) regular classroom instruction, (b) regular classroom instruction but with a portion of the day replaced with one-on-one human tutoring, or (c) regular classroom instruction but with a portion of the school day replaced with computer tutoring on Project LISTEN's Reading Tutor. Mostow et al. [21] describe the study design in full.

To gather results comparable to other published studies on reading instruction, Project LISTEN used the Woodcock Reading Mastery Test (WRMT) [3], an individually administered reading test normed by month within grade to have a mean of 100 and a standard deviation of 15. The WRMT consists of several subtests, each of which tests a specific area of reading skill. In this study, trained testers pre- and post-tested students using the following subtests of the Woodcock Reading Mastery Test: Word Attack (decoding skills), Word Identification (reading single words out loud), Word Comprehension (single word understanding), and Passage Comprehension (understanding 1-2 sentence passages). The testers also measured students' oral reading fluency: their unassisted oral reading rate on prespecified passages at grade level and at a student-appropriate level.

3. Results

Because this paper focuses on vocabulary learning, we discuss here only the results for the Word Comprehension subtest. We analyzed Grade 2 and Grade 3 separately. Figures 1 and 2 show boxplots for pretest, post-test, and gains for normed Word Comprehension scores in Grades 2 and 3, respectively. Figures 4 and 5 show scatterplots of gain by pretest for normed Word Comprehension in Grades 2 and 3, respectively.

We addressed several questions, as follows. Did the children working with the Reading Tutor:

1. gain from pre- to post-test?
2. gain more than a national cohort?
3. gain more than their peers who received classroom instruction?
4. gain more than their peers who received one-on-one human tutoring?

We look at each question in turn.

3.1. Did Reading Tutor students gain from pre- to post-test?

Yes: the difference between post-test and pre-test “raw” weighted score (prior to norming) on Word Comprehension was 15.72 $\pm$ standard error 1.12, with 95% confidence interval (13.49, 17.96). This gain is not all that interesting because it might simply reflect children’s general growth over the year; to filter out such general growth, we next compared to the national norm.

3.2. Did Reading Tutor students gain more than a national cohort?

To answer this question, we looked at the normed gains. (A gain of zero on a normed score means that a student stayed at the same level from pre-test to post-test relative to the norming sample – not that he or she learned nothing, but that he or she learned enough to stay at the same level with respect to the norms.) Students who read with the Reading Tutor averaged gains of 4.38 $\pm$ standard error 0.90, with 95% confidence interval (2.58, 6.18). Therefore students who read with the Reading Tutor learned enough to move forward with respect to the normed scores.

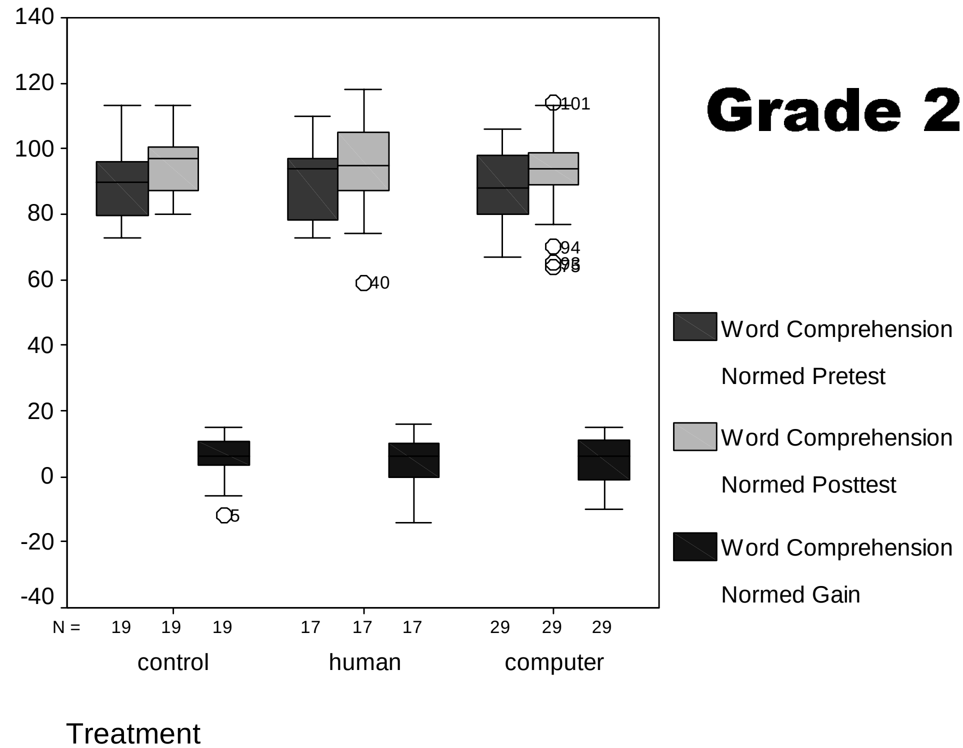

Figure 1. Word Comprehension normed pretest, post-test, and gains, Grade 2. Gain is post-test - pretest.

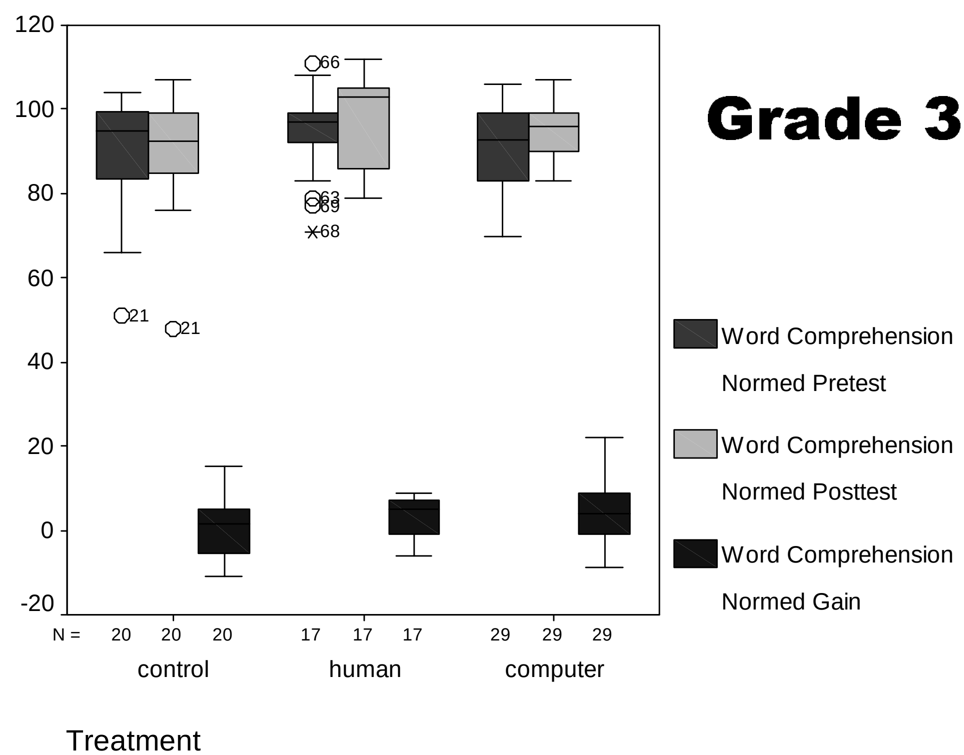

Figure 2. Word Comprehension normed pretest, posttest, and gains, Grade 3. Gain is post-test - pretest.

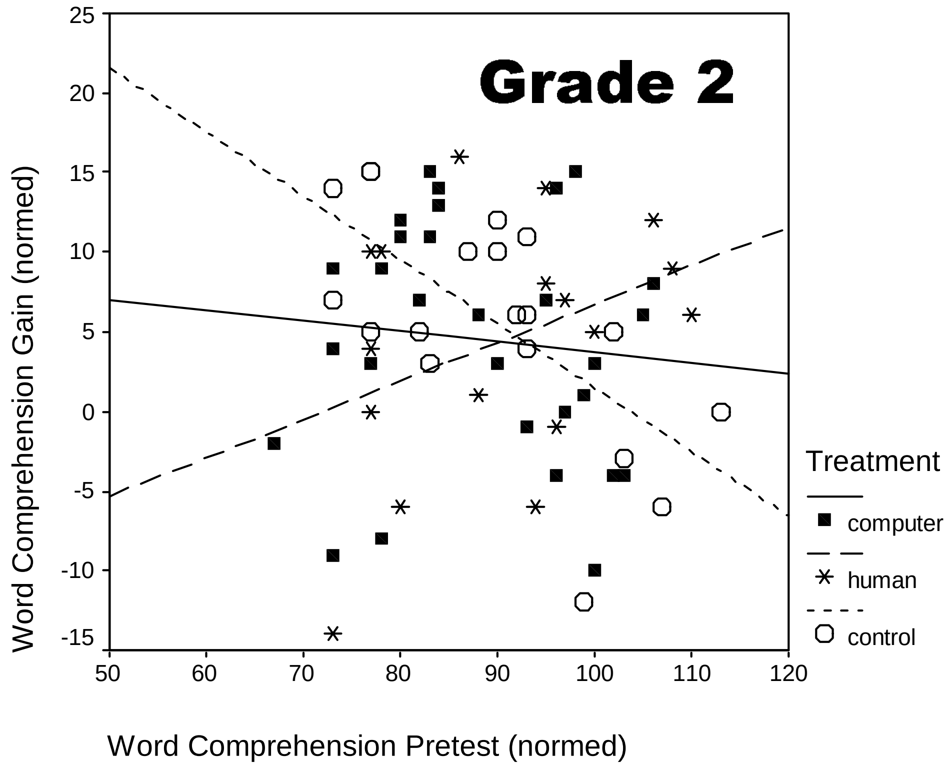

Figure 3. Word Comprehension gains by pretest (normed), Grade 2, with lines showing linear regression fit for each treatment group.

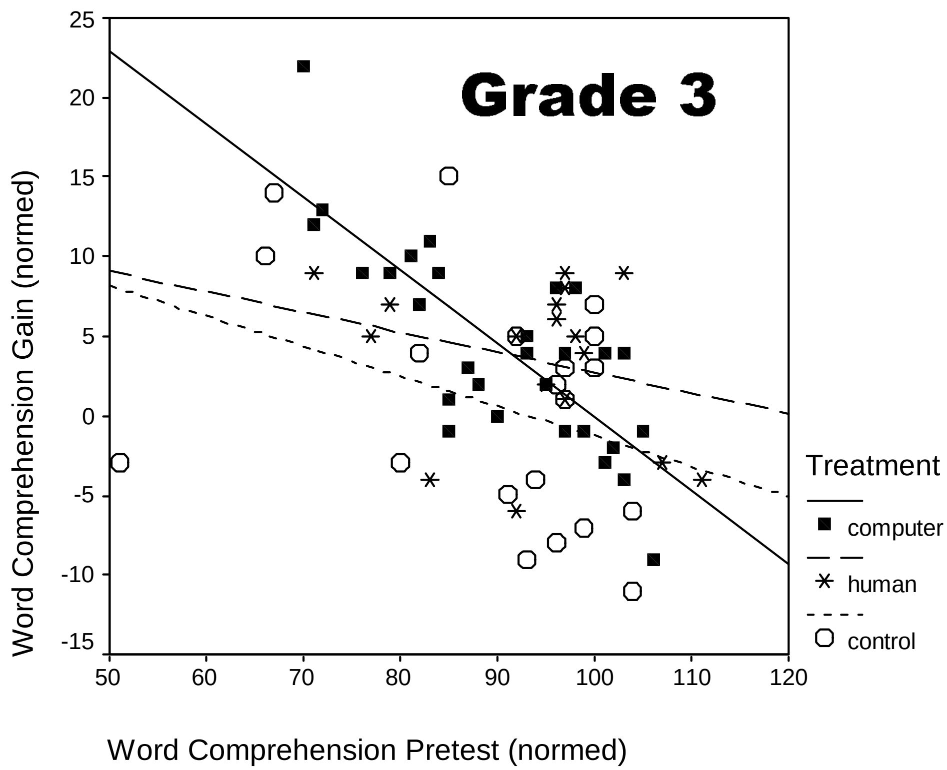

Figure 4. Word Comprehension gains by pretest (normed), Grade 3, with lines showing linear regression fit for each treatment group.

Table 1. Word Comprehension normed score gains by grade, for classroom control (CL), human tutors (HT), and the Reading Tutor (RT). Results prior to covarying out pretest show mean plus or minus standard deviation; results after covarying out Word Comprehension and Word Identification show mean plus or minus standard error.

	Grade 2				Grade 3			
Individual results normed by grade	9.67	± 2.42	HT Room 205	n=6	9.50	± 6.36	CL Room 303	n=2
	7.83	± 6.31	CL Room 208	n=6	7.60	± 5.13	CL Room 305	n=5
	7.00	± 2.65	CL Room 209	n=3	**6.00**	**± 5.18**	**RT Room 301**	**n=8**
	6.17	± 6.71	HT Room 209	n=6	4.80	± 5.17	HT Room 305	n=5
	5.70	**± 7.45**	**RT Room 212**	**n=10**	**4.36**	**± 7.23**	**RT Room 304**	**n=11**
	5.00	± 0.00	CL Room 212	n=2	3.83	± 4.31	HT Room 309	n=6
	4.33	± 10.82	CL Room 205	n=6	**2.90**	**± 6.38**	**RT Room 303**	**n=10**
	3.90	**± 7.92**	**RT Room 211**	**n=10**	2.20	± 6.76	HT Room 310	n=5
	3.67	**± 7.68**	**RT Room 201**	**n=9**	1.00	± 5.66	CL Room 301	n=2
	1.50	± 6.36	CL Room 201	n=2	-4.00	± 4.98	CL Room 310	n=6
	-3.50	± 6.66	HT Room 208	n=6	-4.40	± 5.08	CL Room 309	n=5
Individual results normed by grade and including Word Comprehension pretest and Word Identification pretest as covariates	9.81	± 2.48	HT Room 205		7.80	± 4.33	CL Room 303	
	6.91	± 2.49	CL Room 208		6.29	± 2.74	HT Room 305	
	6.56	± 3.54	CL Room 209		5.92	± 2.75	CL Room 305	
	6.37	± 4.46	CL Room 212		**5.35**	**± 1.85**	**RT Room 304**	
	6.25	± 2.48	HT Room 209		**5.29**	**± 2.16**	**RT Room 301**	
	4.77	**± 1.93**	**RT Room 212**		4.21	± 2.76	HT Room 310	
	4.37	± 2.48	CL Room 205		4.04	± 2.48	HT Room 309	
	3.54	**± 2.04**	**RT Room 201**		**2.46**	**± 1.93**	**RT Room 303**	
	2.76	**± 1.97**	**RT Room 211**		1.57	± 4.30	CL Room 301	
	1.61	± 4.33	CL Room 201		-3.08	± 2.50	CL Room 310	
	-2.22	± 2.52	HT Room 208		-3.96	± 2.73	CL Room 309	

3.3. Did Reading Tutor students gain more than their peers who received classroom instruction?

For the overall results, we used analysis of variance of Word Comprehension gains by treatment and grade, with an interaction term for grade and treatment. To control for regression to the mean, and other differential effects of pretest on gain, we included Word Comprehension pretest. To maximize the fit of the model to the data, we searched through the set of combinations of possible covariates (Word Attack, Word Identification, Word Comprehension, Passage Comprehension, and fluency) and minimized the error remaining between the model and the actual data. This process resulted in using Word Identification pretest as an additional covariate. Analysis of variance using raw scores and considering both grades together revealed a significant interaction between treatment and grade ($F=2.47$, $p=0.088$), so we considered grade 2 and grade 3 separately. Table 1 shows results by room-treatment pairs; since there were six separate human tutors, Table 1 presents results for each human tutor individually, with the Reading Tutor separated out by room for comparison purposes. Table 2 presents an overall summary.

Considering grades 2 and 3 separately, there was no significant effect for treatment in grade 2 ($F=.315$, $p = .731$). In grade 3, there was a significant main effect for treatment ($F=4.27$, $p=0.018$; significant at 95% even if applying Bonferroni correction because there were only two grades in the study). The students on the Reading Tutor in grade 3 did better than their peers receiving classroom treatment, with an estimated advantage on Word Comprehension normed score gains of 3.90 points ± standard error 1.54; for grade 3, effect size[1] = 0.56, and $p = 0.042$, with Bonferroni correction for multiple pairwise comparisons. (Students receiving human

[1] Effect size is the adjusted gains difference divided by the average standard deviation of the compared

tutoring likewise did better than their peers receiving classroom treatment, at 4.56 ± standard error 1.78, effect size = 0.72, p = 0.039 with Bonferroni correction). So, students using the Reading Tutor did better than their peers receiving classroom instruction.[2]

We continued to use this analysis to answer our fourth question.

Table 2. Comparing the Reading Tutor to human tutoring, classroom instruction.

	Overall	Grade 2	Grade 3
Compare classroom-tutored students, human-tutored students, and students who used the Reading Tutor	Marginally significant effect of grade*treatment interaction: F = 2.47, p = 0.088.	No significant main effect of differing treatment conditions	Significant main effect for treatment: F = 4.268, p = 0.018. Human tutor > Classroom: HT – CT = 4.56 ± 1.78 p =0.039 with Bonferroni correction for multiple comparisons Reading Tutor > Classroom: RT – CT = 3.90 ± 1.54 p = 0.042 with Bonferroni correction Reading Tutor ≈ Human tutor: RT – HT = -.663 ± 1.646, p = 1.0
Compare individual human tutors and the Reading Tutor	Marginally significant main effect of different tutors: F=1.8, p = 0.098	Marginally significant main effect of different tutors: F=2.64, p = 0.063	Different tutors not significantly different
Tutors ordered from highest to lowest grade-normed word comprehension gains		Human tutor MB: 9.7 ± 2.4, 6 students Human tutor AC: 6.2 ± 6.7, 6 students Reading Tutor: 4.4 ± 7.5, 29 students Human tutor ME: -4.0 ± 7.3, 5 students	Human tutor LN: 4.8 ± 5.2, 5 students Reading Tutor: 4.3 ± 6.3, 29 students Human tutor MM: 3.1 ± 4.3, 7 students Human tutor NJ: 2.2 ± 6.8, 5 students

3.4. Did Reading Tutor students gain more on Word Comprehension than did their peers who received one-on-one human tutoring?

Again, there was no main effect for treatment in grade 2. In grade 3, there was no significant difference between the human tutored students and the students who read with the Reading Tutor (0.663 points more on normed Word Comprehension gains in favor of the human tutored students, with standard error ± 1.646). How did the Reading Tutor compare to human tutors? Table 1 shows how the Reading Tutor compared to individual human tutors and to classroom instruction on gains for Word Comprehension disaggregated by grade.

Table 2 shows results by grade. For second graders, all three treatment groups achieved approximately equivalent gains. For third graders, the Reading Tutor-ed students gained more than the classroom instruction group on Word Comprehension, and the human-tutored students likewise gained more than the classroom instruction group. How did the Reading Tutor compare to human tutoring? Overall, the Reading Tutor and the human tutors were comparable. Small sample sizes for the human tutors preclude a statistically precise ranking of all the human tutors and the Reading Tutor. Nonetheless, the Word Comprehension gains (both as-is, or adjusted by covariates) by the students in the study placed the Reading Tutor groups roughly interspersed with the human tutored groups.

[2] Incidentally, gains vs. a control group does not permit regression to the mean as an explanation of the

Were there significant differences among the human tutors for Word Comprehension gains? Yes (Table 2). For the second graders, the students tutored by the human tutor M.B. achieved significantly greater gains on Word Comprehension (9.81 ± standard error 2.48) than the students tutored by the human tutor M.E. (-2.22 ± standard error 2.52). Teacher effects, reflected in between-classroom gain differences of as much as 12-15 points in the control condition, may explain some tutor outcome differences more parsimoniously than differences in tutor behavior, given that each tutor covered one room – but not this one: CL 208 >> CL 205, but HT 205 >> HT 208! (Table 2.) To further improve the Reading Tutor on vocabulary learning, we may in the future compare the strategies of the human tutor in Room 205 with those of the human tutor in Room 208.

5. Conclusions

The main result described in this paper is as follows: a computer tutor that did better than classroom instruction for vocabulary learning. The 1999-2000 Reading Tutor even did competitively with one-on-one human tutoring. Mostow et al. [21] give an overview of the entire study. Future work includes reporting this study in detail and investigating the relationship between process variables – how many tutoring sessions students received, the type of interventions that the computer tutor and the human tutor provided, and so forth – and the outcomes of the study.

Acknowledgements

This work was supported in part by the National Science Foundation under Grant Nos. REC-9720348 and REC-9979894, and by Greg 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 thank the anonymous AI-ED reviewers for their comments on this paper.

References

[1] National Institute of Child Health and Human Development, Report of the National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction: Reports of the Subgroups (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office, 2000. Available online from http://www.nationalreadingpanel.org/

[2] C. E. Snow, S. M. Burns, and P. Griffin, Eds., Preventing Reading Difficulties in Young Children. Washington D.C.: National Academy Press, 1998.

[3] American Guidance Service, n.d. Bibliography for Woodcock Reading Mastery Tests – Revised (WRMT-R). http://www.agsnet.com/Bibliography/WRMTRbio.html

[4] M. G. McKeown, I. L. Beck, R. C. Omanson, and C. A. Perfetti, The Effects of Long-term Vocabulary Instruction on Reading Comprehension: A Replication. *Journal of Reading Behavior* **15**, 1 (1983), pp. 3-18.

[5] I. Beck and M. McKeown, Conditions of Vocabulary Acquisition. In *Handbook of Reading Research* vol. 2: pp. 789-814, 1991.

[6] M. G. McKeown, Creating Effective Definitions for Young Word Learners. *Reading Research Quarterly* **28**, 1: pp. 17-31, 1993.

[7] E. B. Zechmeister, A. M. Chronis, W. L. Cull, C. A. D'Anna, and N. A. Healy, Growth of a Functionally Important Lexicon. *Journal of Reading Behavior* **27**, 2: pp. 201-212, 1995.

[8] W. E. Nagy, P. A. Herman, and R. C. Anderson, Learning Words from Context. *Reading Research Quarterly* **20**, 2: pp. 233-253, 1985.

[9] M. G. McKeown, The Acquisition of Word Meaning from Context by Children of High and Low Ability. *Reading Research Quarterly* **20**, 4: 482-496, 1985.

[10] J. P. Gipe, and R. D. Arnold, Teaching Vocabulary through Familiar Associations and Contexts. *Journal of Reading Behavior* **11**, 3: 281-285, 1978.

[11] R. G. Eller, C. C. Pappas, and E. Brown, The Lexical Development of Kindergarteners: Learning from Written Context. *Journal of Reading Behavior* **20**, 1: pp. 5-24, 1988.

[12] J. L. Shefelbine, Student Factors Related to Variability in Learning Word Meanings from Context. *Journal of Reading Behavior* **22**, 1: pp. 71-97, 1990.

[13] R. P. Carver, Ronald P. Percentage of Unknown Vocabulary Words in Text as a Function of the Relative Difficulty of the Text: Implications for Instruction. *Journal of Reading Behavior* **26**, 4: pp. 413-437, 1994.

[14] R. Madden, E. F. Gardner, H. C. Rudman, B. Karlsen, and J. C. Merwin. Stanford Achievement Test. New York: Harcourt, Brace, Jovanovich, Inc, 1973. Cited in (McKeown RRQ 1985).

[15] M. R. Kuhn and S. A. Stahl, Teaching Children to Learn Word Meanings from Context: A Synthesis and Some Questions. *Journal of Literacy Research* **30**, 1: pp. 119-138, 1998.

[16] P. J. Schwanenflugel, S. A. Stahl, and E. L. McFalls, Partial Word Knowledge and Vocabulary Growth during Reading Comprehension. *Journal of Literacy Research* **29**, 4: pp. 531-553, 1997.

[17] D. Reinking, and S. S. Rickman, The Effects of Computer-Mediated Texts on the Vocabulary Learning and Comprehension of Intermediate-Grade Learners. *Journal of Reading Behavior* **22**, 4: pp. 395-411, 1990.

[18] G. Aist and J. Mostow, Improving Story Choice in a Reading Tutor that Listens. Proceedings of the Fifth International Conference on Intelligent Tutoring Systems (ITS'2000), p. 645. Montreal, Canada, June 2000. Poster Abstract. Available online as http://www.cs.cmu.edu/~aist/ITS2000-story-choice-abstract.doc.

[19] G. S. Aist and J. Mostow, Balancing Learner and Tutor Control by Taking Turns: Faster and Better Mixed-Initiative Task Choice in a Reading Tutor that Listens. To appear in Philippe DeCloque and Melissa Holland (Editors), *Speech Technology for Language Learning*. The Netherlands: Swets & Zeitlinger Publishers (in revision).

[20] G. Aist, Factoids: Automatically constructing and administering vocabulary assistance and assessment. Proceedings of the 10th International Conference on Artificial Intelligence in Education (AI-ED), 2001.

[21] J. Mostow, G. Aist, P. Burkhead, A. Corbett, A. Cuneo, S. Eitelman, C. Huang, B. Junker, C. Platz, M. B. Sklar, B. Tobin, A Controlled Evaluation of Computer- versus Human-assisted Oral Reading. Poster presented at the 10th International Conference on Artificial Intelligence in Education (AI-ED), 2001.