Speech Recognition Software Contributes to Reading Development for Young Learners of English

Kenneth Reeder, Department of Language and Literacy Education, University of British Columbia, Vancouver, Canada

Jon Shapiro, Department of Language and Literacy Education, University of British Columbia, Vancouver, Canada

Jane Wakefield, Department of Language and Literacy Education, University of British Columbia, Vancouver, Canada

Reg D'Silva, Department of Language and Literacy Education, University of British Columbia, Vancouver, Canada

ABSTRACT

Thirty-six English language learners aged 6;8 to 12;6 years received practice with The Reading Tutor, which uses speech recognition to listen to oral reading and provides context-sensitive feedback. A crossover research design controlled effects of classroom instruction. The first subgroup worked with the software for 3.5 months, and following a week's crossover period, the second subgroup worked for a subsequent 3.5 months. Both groups were assessed to obtain comparable gains both in regular classroom with English as an Additional Language (EAL) support and in the classroom condition with EAL support plus the Reading Tutor. Oral reading fluency was assessed by the DIBELS measure. Fluency was also calculated by the program, and grade level of materials mastered was assessed by the software's logs. Both groups made significant gains in oral reading fluency and grade level of materials mastered, according to measures internal to the software. For one period, gains in fluency following experience with the program appeared to have been slightly larger than gains with regular classroom instruction and EAL support only.

Keywords: *Automated Speech Recognition, Classroom Instruction, Crossover Research Design, Development, English as an Additional Language, Fluency, Oral Reading, Reading*

DOI: 10.4018/IJCALLT.2015070104

INTRODUCTION

Disproportionate numbers of adult Canadian immigrants are reported to have poor literacy abilities (Statistics Canada, 2005) and this constitutes an obstacle to their acquiring the full range of skills in one of the two official languages of Canada, French or English, as an additional language. A steady increase in the numbers of children with diverse language and cultural backgrounds in hitherto English-dominant centres in Canada mean that well over 50% of the school populations in Vancouver and Toronto have for some years grown up with languages other than English at home (Gunderson & Clarke, 1998), suggesting an increasing proportion of young Canadian learners may potentially experience similar obstacles to acquiring English literacy. This is underlined by results of reading comprehension testing of 41,962 grade 4 students in British Columbia showing that 32% of students designated as learners of English as an Additional Language (henceforth referred to as EAL learners, and their tailored instruction as "EAL support") were reading at levels 'below expectations', compared to 19% of non-EAL students tested (British Columbia Ministry of Education, 2002). At the same time, the long-term educational importance of effective early reading and writing instruction has been long acknowledged (Strickland, 2002).

A useful approach to promoting reading literacy for EAL learners is to customize instruction by integrating technology into classroom teaching (Reeder, Shapiro, Early, Kendrick, & Wakefield, 2008) and leverage the time efficiencies of automation so as to add individualized reading practice to standard group classroom instruction, which has inherent constraints on instructional time that can be devoted to individuals. One limitation of earlier reading software has been its inability to 'hear' readers (D'Silva, 2011; Rasinki, 2013). Recently, prototype software from Project LISTEN, Carnegie Mellon University, entitled The Reading Tutor, (abbreviated to (the) RT throughout) addressed this limitation by using automated speech recognition (ASR) to assist children with oral reading (Mostow, 2001; Project LISTEN, 2007). The RT listens to children read aloud age-graded texts displayed on screen, and offers to read key words or whole sentences aloud, or provide word meanings when children click on a word for help or when the program senses that they are experiencing difficulty as indicated by unusual pauses or otherwise-dysfluent oral performance. Project LISTEN's study with 178 students, Grades 1-4 from schools in the Pittsburgh area found significant fluency and comprehension gains in RT users when the Sustained Silent Reading (SSR) method of instruction was compared to the use of the RT in classrooms (Mostow, Aist, Burkhead, Corbett, Cuneo, Rossbach, et al., 2002). SSR is an instructional method for schools in which all students and teachers devote a set time to leisure reading of their own choice. Studies with EAL learners have also shown similar results. Thirty-four EAL learners in Grades 2-4 whose home language was Spanish in a Chicago suburban school were part of a study that compared SSR instruction in the classroom with the RT and claimed significantly better gains in reading fluency among children using the RT (Poulsen, Wiemer-Hastings, & Allbritton, 2007). Because the RT had not been assessed using young second language learners at the outset of the present project, its potential as an effective literacy intervention for a major population of North American urban children remained unknown. Successful trials of the RT have also been conducted with English language learners in Ghana (Korsah, Mostow, Dias, Sweet, Belousov, Dias, & Gong, 2010) and India (Weber & Bali, 2010).

Technology to Support the Acquisition of Reading

Technologies such as digital audio, Internet technologies and software programs have become popular in the last decade as viable tools in reading instruction in English language contexts. A Statistics Canada report on Information and Communication Technology (ICT) integration in

Canadian elementary and secondary schools claims that governments have recognized the importance of integrating ICT in learning and teaching and have put efforts into installing hardware and software in Canadian secondary schools (Plante & Beattie, 2004). Digital technologies also offer the potential of enhancing oral reading fluency (Rasinski, 2013), which the author notes is a critical element of overall reading proficiency. The National Reading Panel report concluded that "computer technology can be used to deliver a variety of types of reading instruction successfully" (NICHD, 2000, pp.6-9). A number of digital technology tools have been identified by D'Silva (2011) to support classroom fluency and comprehension instruction in meaningful ways and are summarized below.

Simple audio recording of learners' oral reading were found to have improved prosodic performance (i.e., pitch, stress and timing variations) in oral reading (Hudson, Lane, & Pullen, 2005). Talking books and computer-based versions of print books have been assessed as independent reading and repeated reading activities in classrooms. These are merely print books that are augmented with audio recordings. Students could have access to fluency practice "without inordinate demands on classroom teachers and without provoking self-consciousness and frustration on the part of students" according to McKenna, Reinking, and Labbo, (1997, p.185).

Commercially-available interactive multimedia computer software programs are appearing frequently as reading fluency tools. 'Academy of Reading' designed for secondary schools, is proving to be successful in the USA (Loh, 2005). Moreover, reading program packages that consist of a set of graded-readings and activities for the classroom have augmented or replaced their print-based materials with CD- ROM, web-based or software versions. For instance 'Read Naturally', a program that was designed to improve fluency in beginning and older readers with print materials now has a software version (Hasbrouck, Ihnot, & Rogers, 1999; Hudson, et al., 2005). The program claims to use oral reading and repeated reading to promote reading fluency. Interactive singing software technologies like Carry-a-Tune8 (CAT) appeared to be useful in promoting reading skills with 24 struggling readers from grades 7 and 8 (Biggs, Homan, Dedrick, Minick, & Rasinski, 2008).

Automated speech recognition (ASR) software appears to be one of the more promising digital technologies to promote reading proficiency (Rasinski, 2013). In addition to the Reading Tutor (RT) speech recognition software that is reviewed in detail above, the Reading Assistant program, a commercial ASR software program, is designed to improve reading fluency and comprehension in youth and young adults. While its publisher reports upon nine school-based case study trials claiming promising results in reading gains (Scientific Learning, 2015), only one of these (Adams, 2006) was designed as a controlled efficacy study allowing researchers to attribute reported reading gains to the Reading Assistant intervention. Clearly there is a need for further well-designed efficacy studies that can attempt to disentangle the relative impacts upon reading development of real-world variables including traditional classroom instruction, additional instructional programming, and the simple effects of time or learning history, from the specific instructional interventions that are the object of an efficacy study.

Previous Findings of the Vancouver Reading Tutor Project

In previous descriptive studies of the RT, the authors collaborated with colleagues leading Project LISTEN to form the Vancouver Reading Tutor project. We implemented the RT software in schools rich in urban English language learners, contributing three useful discoveries about the appropriateness of the RT for multilingual English language learners.

First, Reeder, Shapiro and Wakefield, (2007, August) and Reeder *et al*., (2008) in a study of 77 participants aged 7-12 years old, included 14 children from Hindi/Urdu speaking house-

holds, 21 Mandarin, 21 Spanish, and 21 native English speaking children 11 of whom completed the RT experience and 10 of whom served as a comparison group. The aim was to determine whether home language background played a role in determining any reading gains yielded by 20 minutes' daily experience on the RT over a 10-week period. All children had been designated to also receive regular EAL pullout instruction, and had been assessed by their school district on a locally developed measure as falling in its 'low intermediate' to 'high intermediate' English proficiency levels. Woodcock Reading Mastery Tests-Revised with the Normative Update (Woodcock, 1998) provided that study's dependent measures of reading, and they were administered in a pretest-posttest design. Analyses of covariance (controlling for age) of mean gain scores for four reading subtests (Word Attack, Word Identification, Word Comprehension, Passage Comprehension) by home language groups showed robust main effects of time for all four home language groups, indicating that participants taken as a whole made significant gains on all four measures over the 10-week period of treatment. The Spanish home language group initially appeared to have made relatively lower mean gains than the three other home language groups on all four reading measures, but that finding was discounted when it was discovered upon closer investigation that the majority of the Spanish speaking participants were recent refugee claimants whose educational history had been badly disrupted over the preceding years to the study. Teachers' reports to that effect are also borne out in a large-scale independent study of immigrant learners in the school district (Gunderson, 2006).

Second, the same study revealed that the lowest of three English language proficiency groups in the sample made greater gains than the higher proficiency groups and native speakers, outgaining the three other groups on Word Identification and Passage Comprehension measures. Their gains compared favourably to those of children in the sample who spoke English as their native language.

Third, in an analysis of the effectiveness of 20 minutes per day with the RT when compared to 30 minutes or more per day of human tutoring over the same six-week span, the two small groups (n=11 and n=10 in RT and Volunteer Tutor groups respectively) of native English-speaking participants were compared on the four reading outcome measures (Reeder, Shapiro, & Wakefield, 2009, July). Analyses of covariance of mean gains revealed main effects of time for all measures for both treatment groups, while paired comparisons analyses revealed that the Volunteer Tutoring group's mean gains were superior to those of the RT group for the Word Comprehension subtest. We tentatively concluded that the RT treatment offered some time efficiencies over conventional human tutoring for at least three of the four reading parameters we studied.

It remained to be seen whether a successful intervention like the RT was inherently superior to classroom instruction with EAL support. The present study extends the Vancouver Reading Tutor Project, and is designed to address the related questions:

1. Is the Reading Tutor superior to classroom instruction with EAL support in improving English reading fluency for EAL learners? and more specifically;
2. In the context of the types of urban schools that this project took place at, are any gains in reading fluency in EAL learners using the RT with EAL support superior to any gains of EAL learners taking only classroom instruction with EAL support?

METHOD

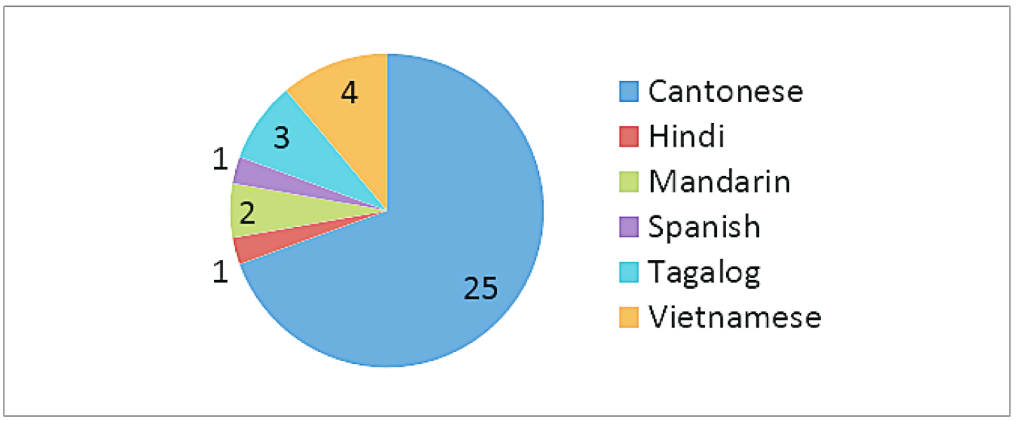
Participants and School Context

Thirty-six students were identified by classroom teachers to take part, (14 female, 22 male,) from one medium sized public elementary school serving a low income, multilingual neighbourhood in Vancouver, Canada. Children were drawn from Grades 2-7, and ranged in age from 6;8-12;6 years. All participants were receiving what we judged to have been expert EAL pullout support for approximately 45 minutes several days per week from a highly experienced EAL specialist teacher, and had been provincially designated for funding for that purpose. The EAL program mandated in the province of British Columbia is designed to transition learners into the mainstream curriculum, which is delivered in English. Participants' English proficiency on the Woodcock Muñoz test ranged from 17 – 33, placing them nominally as beginners to low intermediates in English proficiency at the outset. Children's home languages were roughly representative of the school's catchment area in which the vast majority of households used a language other than English for everyday communication (Figure 1). None of the children's homes reported using English as their principal language.

Dependent Measures

As our primary measure of reading proficiency, we decided to use a widely employed curriculum-based assessment of reading, the DIBELS Oral Reading Fluency (ORF) 6th Edition. Children read a standardized set of passages prescribed for their grade level for one minute. Words omitted, replaced, as well as hesitations for three seconds or longer are scored as errors, while self-corrections within three seconds are accepted as accurately read. The number of words read correctly per minute yields the test score. Its publishers, The Center on Teaching and Learning, University of Oregon, claim that DIBELS assesses "accuracy and fluency with connected text" (Center on Teaching and Learning, University of Oregon, 2015, "Description of the ORF and RTF Measures," para 1). This external measure was not dependent upon administration of the RT and therefore could be administered at the beginning and end of each of the two treatment periods scheduled into the study's design.

Figure 1. Home languages of the study's 36 participants



As secondary measures of reading proficiency, we employed two measures that are embedded into the logging structure of the RT: reading fluency and grade levels of reading passages successfully read. The embedded reading fluency measure counted words recognized by the RT's speech analyzer per minute, averaged over one month intervals, thus it had the advantage of being a repeated measurement. It had the disadvantage that it was administered by an intelligent system that was itself in the process of learning each user's vocal patterns, particularly for the earlier measurements, and hence might have been subject to systematic underestimation of early reading fluency and consequent overestimation of gains over time. The grade level assessments were generated by the RT's internal routine for assignment of reading passage to users dependent upon the internally-measured fluency scores, so they were not strictly independent of that fluency measure. We report gains in grade level of materials mastered from the beginning to the end of the RT treatment periods.

The Crossover Research Design

Crossover research designs (Mills, Chan, Wu, Vail, Guyatt, & Altman, 2009) use participants as their own controls longitudinally, as illustrated in Table 1. Instead of structuring a control group and a treatment group in a two-group design, the design assigns each of two participant groups to a treatment condition as well as a control condition, comparing groups' performance under both conditions. The crossover design allowed us to look at the potential impact of the RT intervention twice, and offered the advantage of adding some control over the simple impact of time, while not denying any participants an educational experience that had independently been shown to be promising and associated with gains in reading proficiency. The design also allowed us to conduct both between- and within-group comparisons. Note that testing occurs at the beginning and end of each of the two 3.5-month treatment periods for both groups in the DIBELS Oral Reading Fluency outcome measure but only before and after the RT treatment (not regular instruction) for those measures that were internal to the RT itself, i.e., fluency with RT materials, grade level of RT materials selected. Hence it is the findings for the DIBELS Oral Reading Fluency measure that would allow us to test more rigorously the efficacy of the RT in comparison to regular classroom instruction with EAL support.

Table 1 also describes the assignment of the study's two subgroups to treatment and control conditions across the study's two 3.5-month treatment periods over the school year, and illustrates the crossover period of one week's duration in mid-February of the year at which point each group would switch treatment conditions. The standardized DIBELS Oral Reading Fluency

Table 1. The study's crossover research design

October	Treatment (3.5 Months)	February	Treatment (3.5 Months)	June
Group 1	Classroom + RT		Classroom Only	
START		CROSSOVER		FINISH
Group 2	Classroom Only		Classroom + RT	

measure was administered to all participants (regardless of their treatment assignment) at the beginning of October, in February during the crossover week, and in June at the conclusion of the study. Data from the two internal reading measures (fluency and grade level) were harvested from the RT software's logs with the cooperation of Project LISTEN at the beginning and end of each of the two RT treatment phases only for the group assigned to the RT. Gain scores were calculated for all measures.

Assignment to the two subgroups required for the crossover design was carried out by the participants' own classroom teachers, using timetabling constraints as the main consideration. Table 2 illustrates the composition of each subgroup in terms of some relevant variables that might bear upon reading development. In terms of similarity, groups came close to providing us with a matched sampling design for the study, but children were teacher-assigned, not randomly assigned, to groups, since this was an ordinary school with fixed timetables and the usual constraints rather than a laboratory. Thus, the design is quasi-experimental, not a fully randomized experiment. That limitation was, we believe, offset to a degree by the ecological validity of using intact school groups in more naturalistic groupings and normal timetabling.

Procedure

Two Reading Tutor systems were installed centrally in the school library under a Teacher Librarian's supervision (see Figure 2), who volunteered to deal with startup and shutdown at beginning and end of each school day, and to alert the research team of any technical difficulties encountered in the daily operation of the program or the hardware. Teachers worked with the research team to schedule each child for 20 minutes' practice with the software daily for the 3.5 months in either the first or second treatment period according to the design.

The investigators' RT treatment objective was 20 minutes' daily practice, over 4 days per week on The Reading Tutor for just over 3 months. The actual values achieved by participants slightly exceeded this objective and demonstrated good standardization of treatment on RT given the similarity of their profiles of participation:

- Group 1 mean (range) participation:
  - 52.0 days (min 40 days, max 61 days);
  - 19.9 hours (min 14.0 hours, max 29.8 hours);
- Group 2 mean (range) participation:
  - 51.3 days (min 32 days, max 64 days);
  - 19.3 hours (min 11.0 hours, max 28.3 hours).

Table 2. Composition of the two subgroups used in the study's crossover design

	GROUP 1, n=18	**GROUP 2, n=18**
Age on entry	8.7 yrs (6;8-11;4)	8.6 yrs (7;0 – 12;6)
Gender	F=6 M=12	F=8 M=10
English Proficiency	26.7 (19-33)	25.0 (17-30)
Initial Reading fluency	60.3 words/min (16-108)	51.7 words/min (9-120)
Initial Reading comprehension	23.5 (6-40)	23.2 (6-37)

Figure 2. The study's configuration of The Reading Tutor in its school setting

Photo credit: Lei Hong. Used with permission.

The working hypothesis for our comparison of the two treatment conditions was that if the RT treatment was indeed more effective in promoting oral reading fluency than classroom experience with EAL pullout support alone, then marginal mean reading fluency gain scores should resemble the idealization illustrated in Figure 3. In that idealization, students experiencing classroom teaching plus the RT would outperform students experiencing only classroom teaching, and further, each group would attain higher reading fluency gains under the RT treatment condition than under the classroom only condition. The idealized performance curves would thus resemble a diamond-shaped figure.

FINDINGS

We begin by presenting results gained from the two internal measures we accessed from Reading Tutor's user logs, its reading fluency measure, and its grade level measures, recorded at the start and finish of each group's assignment to the RT treatment. We then turn to the findings from the DIBELS Oral Reading Fluency test which was administered before and after all treatment conditions, offering us the possibility of comparing the efficacy of the two conditions in promoting oral reading fluency.

Figure 4 displays the mean gains in internally-assessed reading fluency from start to finish of the RT treatment for each group. We found a main effect of time ($F=4.92$ (1,34), $p=.03$) with an effect size of .126. We found no between group effect, in the sense of order of treatment advantage (fall vs. spring) for this measure, indicating that all students made strong gains in reading fluency on this measure, regardless of treatment order.

Figure 5 presents mean grade level gains with RT materials, assessed at the start and finish of each RT treatment phase. Again, there was a main effect of time found ($F=72.55$ (1,34), $p=.000$)

Figure 3. Idealized curves assuming greater efficacy of RT treatment condition

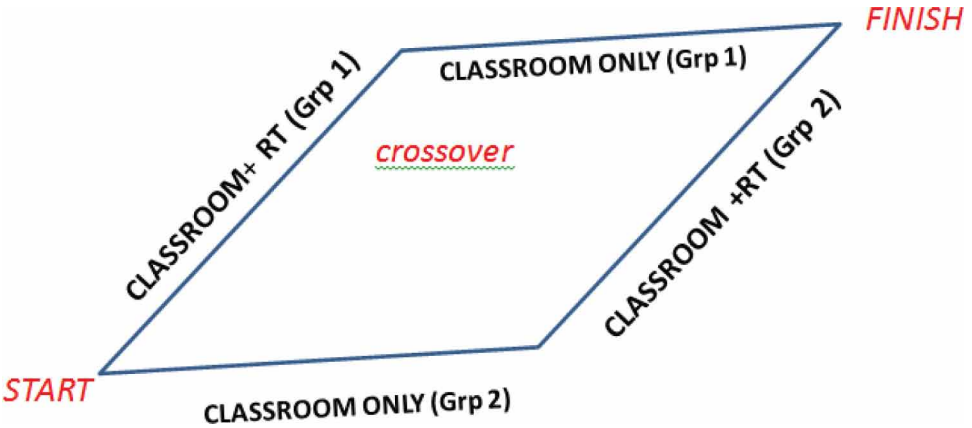
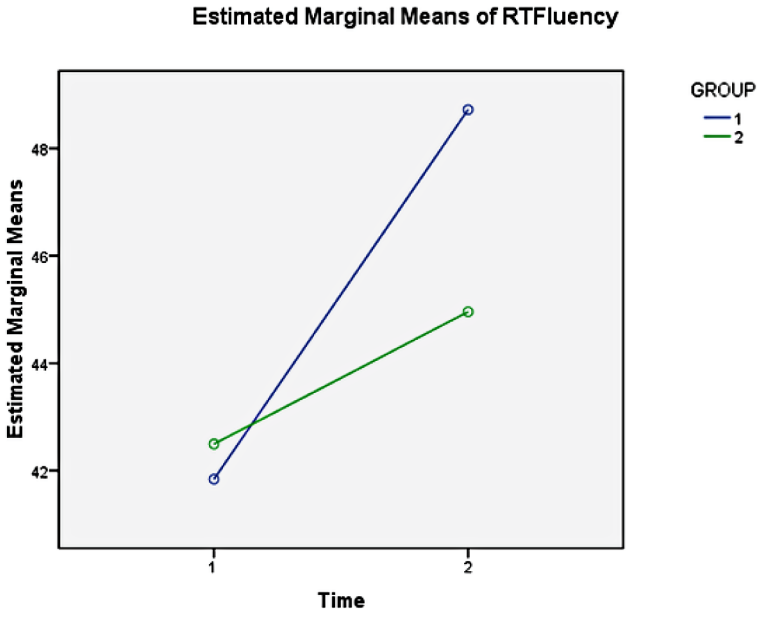


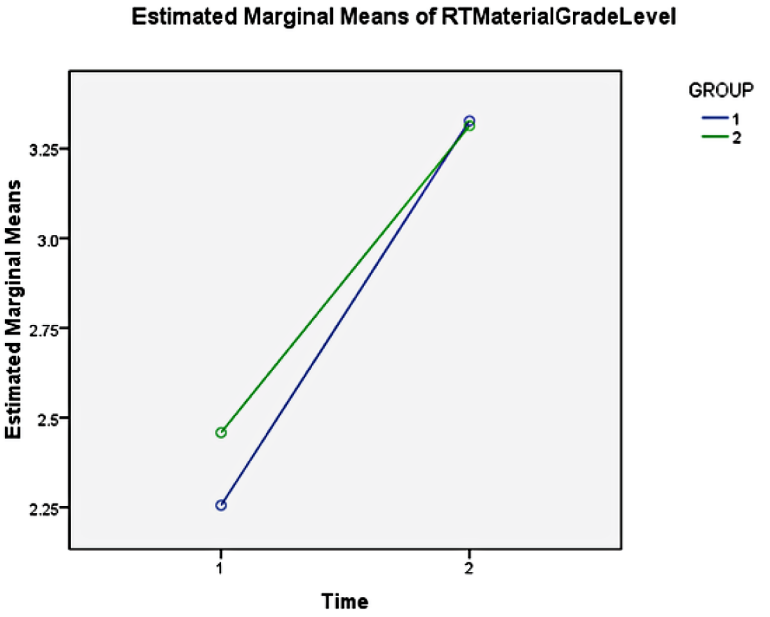
Figure 4. Internally measured fluency gains, start to finish of each RT treatment



with a substantial effect size of .681. And again, there was no group effect, in the sense that there appeared to have been no treatment order advantage, suggesting that all students made strong gains in grade level of materials fluently read, roughly a one-year's gain over the 3.5 months of treatment, regardless of treatment order.

These findings derived only for the RT intervention are naturally descriptive only, since there was a confounding variable during the RT treatment periods, i.e., regular classroom experience together with expert EAL support. Would our standardized assessment of oral reading fluency administered not only before and after the RT treatments, but also before and after the

Figure 5. Mean grade level gains from start to finish of each RT treatment period



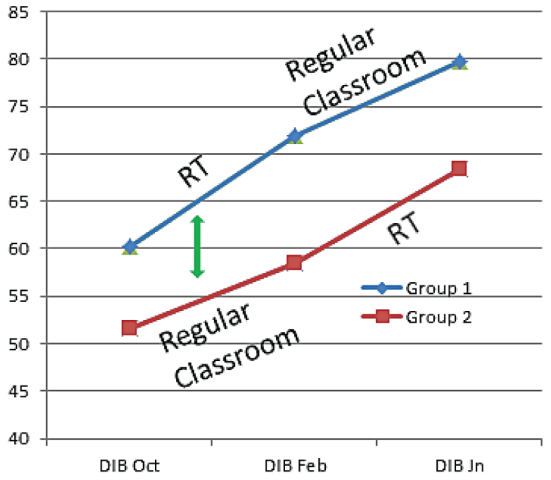
classroom conditions provide comparative findings that might shed light upon the efficacy of the Reading Tutor treatment?

Figure 6 compares mean gains on the standardized DIBELS Oral Reading Fluency measure for the two subgroups before and after both treatment conditions. There was a substantial main effect of time, (F=33.23 (2,66), *p*=.000) with an effect size of .412. We also found a potential but non-significant effect of group, in which Group 1's mean gain score was slightly superior to that of Group 2 (F=1.195 (1,33), *p*=.282) with a very small effect size, .035. Similarly, we found a potential but non-significant treatment difference, suggesting that the RT group's mean gains were slightly superior to those of the classroom treatment group's in the October through February phase of the study (F=1.47 (1, 34), p=.23) again with a very small effect size, .043. This is noted by means of a vertical green arrow superimposed upon Figure 6, since this is the only suggestion in the findings that the RT in itself is potentially more effective than regular classroom instruction at enabling students to increase their oral reading fluency. While Figure 6 does indeed illustrate curves that move somewhat in the direction of Figure 3's idealized confirmation of the RT's efficacy when tested under the realistic conditions of good classroom instruction together with high quality EAL pullout support for English language learners, only in one of this study's two parallel trials was there a suggestion of the RT's superiority in promoting oral reading fluency, and that indicated a very small effect size. And this is not to diminish the role that the RT very probably played in contributing to the substantial reading gains achieved by both groups in our study, as we discuss below.

DISCUSSION AND LIMITATIONS OF THE STUDY

We need to consider why those fluency increases specifically associated with the use of The Reading Tutor were so modest (effect sizes .035 and .043) despite the impressive gains made

Figure 6. Mean gains on the DIBELS oral reading fluency measure for the two subgroups before and after treatment conditions



by all students over the school year (effect size .412). We believe there are three reasons for these findings, and these relate to inherent limitations posed by the study's design and field implementation.

Effect of Short Crossover Period on Phase 2 Findings

It is conceivable that the oral reading fluency improvement associated with Group 1's experience on The Reading Tutor was cumulative, and continued to affect reading after the very short crossover or 'washout' period, only one week in length. Those early gains probably contributed to subsequent performance under regular classroom conditions without the RT, decreasing the within-group contrast by treatment for that group. Mills et al. (2009) note that crossover designs are most appropriate for studies whose treatment effects are short-lived and reversible. Clearly, early reading practice is not such an experience but is regarded by educators as continuing in its benefits, providing a foundation for later learning across the curriculum (Strickland, 2002).

Moderating Effect of Good Quality Classroom Instruction with Strong EAL Support for All Groups

This study took place in a well-run, reasonably well-funded school system that made a priority of supporting its majority population, who were English language learners from lower-income, multilingual homes. The investigators observed a good level of staff morale and experienced enthusiastic cooperation from teachers and school leadership alike. It might well have been that the large, main effects of time over the school year were exactly the findings that an independent assessment of reading development should have found, in the light of the relatively short exposure to the sophisticated ASR program that the RT represents: its impact was simply overwhelmed by good teaching and high quality support for English Language Learners. Quality classroom instruction – in its absence - may well have been the reason for the very promising results for a much shorter crossover classroom trial of the RT found by Poulsen, Wiemer-Hastings, and Allbritton (2007). That study reported difficulties at the school level in recruitment and coopera-

tion, which suggested that its school setting was less than optimum, whereas the present study was carried out under optimal instructional conditions in our view.

Did the RT Contribute to Participants' English Language Development?

It is arguable, to the extent that reading is a language process rather than a disembodied cognitive skill, that the RT could have made a positive contribution to the fundamental English Language learning that all participants were undergoing during the school year. This is because the RT appears to promote one of the essential elements of good English language instruction itself. Second language acquisition research tells us that 'noticing errors' is crucial in promoting language error correction and learning more generally (Schmidt, 2012). Does the feedback provided by the RT promote language development in that respect? We would argue that the RT promotes noticing effectively because of three key design features:

- The feedback provided by the RT is subtle: It merely does not highlight incorrectly read text as the reader proceeds; further, it signals that the program is waiting for improved input only if there is a major delay or error;
- Feedback provided by the RT is imperfect in the sense that its ASR technology does not catch every single reading error and stop to offer help (Mostow & Aist, 1999). Consequently, discourse flow is not frequently interrupted, thus helping to balance contending fluency and accuracy performance goals. The RT offers more human-sounding, less robotic discourse interaction;
- Because feedback from the RT occurs in private, with no public failures or distractions, noticing an error is less stressful and potentially more easily undertaken.

For these reasons, we believe that the RT, for all its demonstrated benefits, did not differentiate itself sufficiently from the strong baseline of classroom instruction and language support that was in place in the present study for it to contrast greatly in its contribution to oral reading fluency development.

CONCLUSION

This crossover-design study was able to demonstrate gains in oral reading fluency for English language learners while undertaking several months' practice with The Reading Tutor. The first measure, internal to the RT software's logging system, showed significant gains for participants from beginning to end of their 3.5 months of RT use with an effect size of .126. The second measure, also internal to the RT's software, showed that participants by the end of their 3.5 months of RT practice were tackling reading materials one full year's reading level in advance of the materials they had begun with, showing a substantial effect size of .681. An independent, standardized procedure (DIBELS) for assessing gains in oral reading fluency revealed fluency gains following the RT treatment that were slightly larger than gains following regular classroom instruction and EAL support only for one of the administration periods of the study, the fall term. Although the effect sizes for the two group's gains under the RT condition, .035 and .043 were modest, we found a very strong overall gain in fluency for the two groups combined across the school year, with a substantial effect size of .412. In terms of the study's related research questions that concerned the superiority of the RT experience over the benefits of good classroom instruc-

tion coupled with expert English language support, our findings should be treated as cautious affirmations of our working hypothesis, given the modest effect sizes of our comparative data.

It is nonetheless safe to conclude that the RT was clearly associated with the observed strong gains in oral reading fluency when measured independently or within the intervention's own assessment tools, and with significant grade level gains in passages mastered. As such, the present study showed that it is a worthwhile addition to even optimal instruction and language support, and that it may indeed have contributed to the strong main effects of time our study revealed. Only replications of this work, ideally with a larger sample size and longer treatment periods, potentially more robust approaches to the assessment of reading including comprehension (cf. Fuchs, Fuchs, Hosp, & Jenkins, 2001) and perhaps even deployment in less optimal school settings can resolve more definitively the question of the relative efficacy of the RT in relation to 'regular' classroom teaching.

ACKNOWLEDGMENT

An earlier version of this paper was presented as: Reeder, Shapiro, & Wakefield (2014, August). The authors wish to thank Lei Hong for her contribution to data collection for the study reported here, and acknowledge the kind support of the Vancouver School Board, and the Social Sciences & Humanities Research Council, Canada, in its grant support to Kenneth Reeder and co-investigators.

REFERENCES

Adams, M. J. (2006). The promise of automatic speech recognition for fostering literacy growth in children and adults. In M. C. McKenna, L. D. Labbo, R. D. Kieffer, & D. Reinking (Eds.), *International Handbook of Literacy and Technology* (Vol. 2, pp. 109–128). Mahwah, NJ: Erlbaum.

Biggs, M. C., Homan, S. P., Dedrick, R., Minick, V., & Rasinski, T. (2008). Using an interactive singing software program: A comparative study of struggling middle school readers. *Reading Psychology*, *29*(3), 195–213. doi:10.1080/02702710802073438

British Columbia Ministry of Education. (2002). BC performance standards: Reading. Victoria, BC, Canada.

Center on Teaching and Learning, University of Oregon. (2015). *DIBELS Data System*. Retrieved fromhttps://dibels.uoregon.edu/market/assessment/measures/orf.php

D'Silva, R. A. (2011). *Promoting reading skills of young adult EAL learners through voice recognition software.* Unpublished doctoral dissertation, The University of British Columbia, Vancouver, BC.

Fuchs, L., Fuchs, D., Hosp, M., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256. doi:10.1207/S1532799XSSR0503_3

Gunderson, L. (2006). *English-only instruction and immigrant studies in secondary schools: A critical examination*. Mahwah, NJ: Erlbaum.

Gunderson, L., & Clarke, D. (1998). An exploration of the relationship between ESL students' backgrounds and their English and academic achievement. In T.Shanahan, F. V.Rodriguez-Brown, C.Worthman, J. C.Burnison, & A.Cheung (Eds.), 47th yearbook of the National Reading Conference (pp. 264–273). Chicago, IL: National Reading Conference.

Hasbrouck, J. E., Ihnot, C., & Rogers, G. H. (1999). Read Naturally: A strategy to increase oral reading fluency. *Reading Research and Instruction*, *39*(1), 27–37. doi:10.1080/19388079909558310

Hudson, R., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher*, *58*(8), 702–714. doi:10.1598/RT.58.8.1

Korsah, G. A., Mostow, J., Dias, M. B., Sweet, T. M., Belousov, S. M., Dias, M. F., & Gong, H. (2010). Improving Child Literacy in Africa: Experiments with an automated reading tutor. *Information Technologies and International Development*, *6*(2), 1–19.

Loh, E. (2005). *Building reading proficiency in high school students: Examining the effectiveness of the academy of READING for striving readers.* Retrieved January 31, 2008, from http://www.autoskill.com/pdf/HS_metastudy2005.pdf

McKenna, M., Reinking, D., & Labbo, L. D. (1997). Using talking books with reading-disabled students. *Reading & Writing Quarterly*, *13*(2), 185–190. doi:10.1080/1057356970130206

Mills, E., Chan, A., Wu, P., Vail, A., Guyatt, G., & Altman, D. (2009). Design, analysis, and presentation of crossover trials. *Trials*, *10*(1), 27. doi:10.1186/1745-6215-10-27 PMID:19405975

Mostow, J., & Aist, G. (1999). Giving help and praise in a reading tutor with imperfect listening – because automated speech recognition means never being able to say you're certain. *CALICO Journal*, *16*(3), 407–424.

Mostow, J., & Aist, G. (2001). Evaluating tutors that listen: An overview of Project LISTEN. In K. Forbus & P. Feltovich (Eds.), *Smart Machines in Education* (pp. 169–234). Palo Alto, CA: MIT/AAAI Press.

Mostow, J., Aist, G., Burkhead, P., Corbett, A., Cuneo, A., Rossbach, S., ... (2002). *Independent practice versus computer-guided oral reading: Equal-time comparison of sustained silent reading to an automated reading tutor that listens.* Paper presented at the Ninth Annual Meeting of the Society for the Scientific Study of Reading, June 27-30, Chicago, IL.

NICHD. (2000). *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: Report on the subgroups (No. NIH Pub. No. 00-4754)*. Rockville, MD: National Institute of Child Health and Human Development.

Plante, J., & Beattie, D. (2004). *First results from the Information and Communications Technologies in schools survey, 2003-2004.* Retrieved from http://www.statcan.ca/english/research/81-595-MIE/81-595-MIE2004017.pdf

Poulsen, R., Wiemer-Hastings, P., & Allbritton, D. (2007). Tutoring bilingual students with an automated reading tutor that listens. *Journal of Educational Computing Research*, *36*(2), 191–221. doi:10.2190/A007-367T-5474-8383

Project Listen. (2007). *Project Listen Summary*. Retrieved from http://www-2.cs.cmu.edu/~listen/

Rasinski, T. (2013). *Supportive fluency instruction: The key to reading success (especially for students who struggle). A white paper for Scientific Learning.* Scientific Learning, Oakland, CA. Retrieved from http://www.scilearn.com

Reeder, K., Shapiro, J., Early, M., Kendrick, M., & Wakefield, J. (2008). Listening to diverse learners: The effectiveness and appropriateness of a computer-based reading tutor for young Canadian language learners. In F. Zhang & B. Barber (Eds.), *Handbook of research on computer-enhanced language learning* (pp. 159–188). Hershey, PA: IGI. doi:10.4018/978-1-59904-895-6.ch010

Reeder, K., Shapiro, J., & Wakefield, J. (2007, August). The effectiveness of speech recognition technology in promoting reading proficiency and attitudes for Canadian immigrant children. Paper presented at the 15th European Conference on Reading, Berlin, Germany.

Reeder, K., Shapiro, J., & Wakefield, J. (2009, July). A computer based reading tutor for young English language learners: Recent research on proficiency gains and affective response. Paper presented at the 16th European Conference on Reading and 1st Ibero-American Forum on Literacies, Braga, Portugal.

Reeder, K., Shapiro, J., & Wakefield, J. (2014, August). *Advanced speech recognition supports reading development for young EAL learners.* Paper presented at the World Congress 2014, International Association of Applied Linguistics, Brisbane, Australia.

Schmidt, R. (2012). Attention, awareness, and individual differences in foreign language learning. In W. Chan, K. Chin, G. Bhatt, & I. Walker (Eds.), *Perspectives on individual characteristics and foreign language education* (pp. 27–50). Boston, MA: Walter de Gruyter. doi:10.1515/9781614510932.27

Scientific Learning Inc. (2015). *Reading Assistant ™ Results.* Retrieved from http://www.scilearn.com/results/reading-assistant-results

Statistics Canada. (2005). *Adult Literacy and Life Skills Survey.* Retrieved from http://www.statcan.gc.ca/daily-quotidien/050511/dq050511b-eng.htm

Strickland, D. (2002). The importance of early intervention. In A. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 69–86). Newark, DE: International Reading Association.

Weber, F., & Bali, K. (2010). Enhancing ESL education in India with a reading tutor that listens. In *Proceedings of the First ACM Symposium on Computing for Development* (20: pp. 1–9). New York, NY, USA: ACM. doi.org/ doi:10.1145/1926180.1926205

Woodcock, R. (1998). *Woodcock reading mastery tests – Revised/normative update*. Circle Pines, MN: American Guidance Service.

Kenneth Reeder is Professor in the Department of Language & LIteracy Education at the University of British Columbia, Vancouver, Canada.

Jon Shapiro is a Professor of Literacy Education at the University of British Columbia. His research has focused on early stages of reading development with a particular focus on affective factors.

Jane Wakefield is a doctoral candidate in the Human Development, Learning, and Culture program in the Department of Educational and Counselling Psychology, and Special Education, Faculty of Education,The University of British Columbia.

Reg D'Silva is a 12 month-lecturer in the Language and Literacy Education Department (LLED) at the University of British Columbia, where he has taught undergraduate TESL and language through content courses. His research interests are in TESL, international education and the integration of technology in language and literacy contexts.