A Large Scale Study of the Effectiveness of Multi-Sensory Learning Technology for Learning English as a Second Language

Swini Garimella
Venkat Srinivasan

March 2014

Abstract

There is a large and growing body of literature on the efficacy and positive impact of technology for improving English language learning. The vast majority of these studies have examined the impact of adopting technology such as digital classroom, videos, and mobile platforms on the proficiency of learners. More recently electrophysiological studies of letter–speech sound processing have revealed that multi-sensory learning approaches have significant potential to reduce the effort in mastering the ability to read.

In this study, we provide evidence on the impact of multi-sensory technology for English language learning in the case of non-native learners. We report the results of a large-scale application of multi-sensory English language learning technology on approximately 20,000 students in 100 schools across 6 states in India. The results reveal significant improvement in reading, reading comprehension, vocabulary, grammar and spelling in English language across the board despite the absence of adequate infrastructure and other constraints inhibiting learning.

Introduction

The role of technology as a resource for instruction of foreign language learners is increasing as educators recognize its ability to create both independent and collaborative learning environments in which students can acquire and practice a new language. A large number of studies have documented the impact of different technologies on English language learning [ELL] populations generally concluding that the use of technologies have been beneficial.

In developing economies like India, English proficiency takes on a whole new meaning and social significance. It is tied to economic and social well-being. Several studies document the differential income generating capacity of persons with superior English language skills. English is clearly 'aspirational' in such countries.

We should recognize that 'technology' encompasses a wide range of tools, artifacts and practices, from multimedia computers to the internet, from videotapes to chat rooms, from web pages to interactive audio conferencing and more. Besides, the effectiveness of any specific technology is highly influenced by many other variables – the environment, the teacher, the instructional method and setting, and the assessment approach. Even the use of the same technology across these variables may result in different learning outcomes.

Recent studies from neuroscience suggest that multi-sensory structured learning education [MSLE] has significant potential to improve the systematic acquisition of reading skills. These studies have found that the human brain was not created to recognize the letter-speech sound combinations required for reading fluency. Letter-speech sound variations are arbitrary cultural inventions. The brain creates a specialized neural pathway for recognizing such arbitrary objects. Multi-sensory stimulation enables the speedy creation of such a neural network.

In this study, we report on the results of a large-scale application of a multi-sensory software application for English language learning across 100 government schools in India covering 20,000+ students in grades 6, 7, and 8. The results indicate that MSLE proved to be highly effective in English language learning in highly constrained, infrastructurally challenged settings.

The study is pioneering in many ways. First, it is the first large scale study of the impact of technology for English language learning and specifically reading skills, in a K-12 setting. Second, it supports significant research findings that multisensory methods can be effective in the acquisition of reading skills. Third, the study was conducted in government schools across India in settings that have poor infrastructure, i.e., many students sharing a single computer, limited access to the application, poor classroom facilities, and so on. Fourth, the

students who were part of the study have little or no exposure to English outside the classroom. Fifth, the study period was sufficiently long to enable robust conclusions about the impact of the software application.

Prior Research

In the eighties, the application of technology in language classrooms included the use of film, radio, television, language labs with audio/video tapes, computers, and interactive video (Cunningham, 1998). Various types of computer-assisted language learning (CALL) also began to become more commonplace (Iandoli, 1990). As the technology advanced, we began to see more interactive uses of CALL as well as an increase in the integration of various media into the computer system (Pusack & Otto, 1990). Computer technology became more accessible to both individuals and schools. Moreover, our growing understanding of its potential has encouraged a shift in emphasis from computer technology itself to its applications. Today, the use of multimedia, the Internet (especially the World Wide Web), and various forms of distance learning are wide spread. Interest in using them as tools to support language learning is growing, both from the perspective of a language educator and that of a language learner.

Liu, et al (2001) reviewed the research on computer-based technology use in second language learning during 1990-2000. They sorted research articles into the following categories: (i) non-research based, and (ii) research based. The non-research based category included conceptual discussions and project descriptions on (a) the potentials of computer technology and its use in specific areas, (b) software tools used in certain language skill areas, (c) software design considerations, and (d) computerized language testing. They found that findings from numerous studies suggested that the use of visual media supported vocabulary acquisition and reading comprehension, and helped increase achievement scores. They also found that more research needs to be conducted at the K-12 level.

Ban et al (2000) review research on the use of technology for English language learners specifically in Content-based classrooms. They argue that it is no longer the question of whether computer technologies have the potential to help ELL students develop English language proficiency. The outstanding issue according to them is how to effectively integrate technology into content-based classroom pedagogy.

Many other researchers have attempted to provide a comprehensive review of studies documenting the impact of technology on language learning [e.g., Cavanaugh, 2001; Chapelle, 1997; Lou, Abrami and d'Apollonia, 2001; Salaberry, 2001]. Zhao points out that there has also been a major paradigm shift in the

pedagogical research focus of technology applications in language education recently [Chapelle, 1997, 2001; Pennington, 1996; Salaberry, 2001] – a shift away from traditional drill-and-skill computer-aided instruction (CAI) models toward multimedia, intelligent CAI, and integration models.

Zhao applied meta-analysis on all the qualified empirical studies included in his review and reported a significantly positive impact of technology applications on language learning [see table below]. For the meta-analysis, more than one effect size was calculated for several studies because they had more than one measure (e.g., listening, reading and writing). In order to satisfy the independence assumption of meta-analysis only one effect size per study was included in the study. The effect sizes used in the analysis were weighted ds, which corrected sample size biases [Hedges & Olkin, 1985].

Table 1
Overall Effect of Technology Applications in Language Learning

Analysis	K	N	Mean Weighted d	Std. Deviation	95% Confidence level for d
Averaged	9	419	+1.12	0.78	0.61 to 0.63
All	29	1045	+0.81	0.72	0.55 to 1.07

As the above table shows, the mean effect size of the 9 studies is quite large, indicating an overwhelmingly positive effect of technology applications on language learning. Zhao did not differentiate among the areas of improvement in the target language. The results were a summary of the empirical findings of the effects of a variety of technological applications on virtually all aspects of language learning – vocabulary, grammar, reading, listening, writing, speaking and cultural understanding.

Zhao (2003) provides some critical insight from the review:

1   First, the total number of well-designed experimental studies on the effectiveness of technological applications in language learning is very limited. Very few of the articles found were empirical studies. Even they did not measure learning gains, relied solely on learner self-reports as measures of effectiveness or were not well designed.

2   Second, the studies were limited to college-level language learners. None of the studies reviewed were conducted in K-12 settings.

Most of the above studies suffer from several significant limitations. First, the sample sizes of all the studies are relatively small and suffer from a selection bias. Second, there are very few studies that focus on K-12 students; most focus on college learners. Zhao (2003) makes the observation that none of the studies

found in the major language education and technology journals is about technology use in K-12 classrooms whereas most such studies in other subject areas (mathematics, science, social studies and language arts) have taken place mainly in K-12 settings. The two populations have very different motivations and at very different stage in their evolution. Third, it seems like in many cases the instructors designed, implemented and evaluated the assessments, which of course potentially makes the results questionable.

Multi-Sensory Methods for Language Acquisition

More recently, there is significant evidence from neuroscience that multi-sensory learning is likely to be more effective for language acquisition instead of uni-sensory approaches. Shams and Seitz (2008) point out that studies of learning and in particular perceptual learning have typically focused on learning of stimuli consisting of a single sensory modality. However, we constantly experience multisensory stimulation in the real world.

It is easier to integrate multiple sources of information during learning when the material is physically integrated, auditorily and visually, than when information is presented to each modality separately [Mousavi, Low & Sweller (1995)]. It appears that multi-sensory information processing is part and parcel of object perception and recognition in daily life, whereby the brain integrates the information from different modalities into a coherent percept [Ghazanfar and Schroeder, 2006]. Therefore, it is likely that the human brain has evolved to develop, learn, and operate optimally in multisensory environments. These studies suggest that multisensory training protocols can better approximate natural settings and are more effective for learning.

Blomert and Froyen (2010) make the point that in the last decade, neuroimaging studies have identified a brain region that shows specialization for fast visual word recognition i.e., the putative Visual Word Form Area (Cohen et al, 2000) in the occipito-temporal cortex. Since fluency and automaticity are the most salient features of experienced reading, it is indeed plausible that a neural network involved in visual object recognition has specialized for recognizing visual letters and word forms (McCandliss et al., 2003).

Years after children first learn to decode letters into words, a form of perceptual expertise emerges in which groups of letters are rapidly and effortlessly conjoined into integrated visual percepts, a process which is crucial to fluent reading ability. We need years of explicit instruction and practice before we start to exhibit any fluency in visual word recognition. This contrasts sharply with the way we learn to master spoken language. Infants and young children start to pick up and develop the many complexities of spoken language without explicit instructions at a time when literacy instruction is still far in the future. Recent

6

electrophysiological evidence showed that it takes several years of reading instruction and practice before the first signs of automatic integration of letters and speech sounds appear in normally developing children. Letter –speech sound associations are cultural interventions and therefore biologically arbitrary in nature.

While MSLE holds much promise to accelerate reading skills, there is little empirical evidence on its efficacy in the real world.

Study Design & Results

In India's government schools, reading and literacy levels are extremely low. The Annual Status of Education Report (ASER), published on the state of education in rural India, declared that nationally, the proportion of children in grade three able to read at least a paragraph of grade one is still abysmally low. In 2013, only two out of five children could achieve this standard. Similarly, the proportion of children in the fifth grade at the all-India level who could read a second grade text remained unchanged at the level of 47%. It has decreased every year from 52.8% in 2009. While enrolment in elementary education has increased and is almost 100%, the bad news is that the education outcomes, as measured by abilities in reading, writing and math, have deteriorated among children between the ages of six and 14.

This study has added significance in the context of India's socio-economic development in the coming decades. In India, English is the language for social empowerment, higher education, business and governance. Millions of young people are emerging academically undernourished from India's education system. By 2020, India has to assist 500 million youth in becoming employable. Poor reading ability and proficiency in English is a critical issue that needs to be solved.

EnglishHelper [EH] is marketing an English language learning application called ReadToMe™ [RTM]. RTM is a multi-sensory application allowing the learner to see, hear and speak the language and gain proficiency. Originally developed as a remedial tool for native English speakers struggling to learn how to read, its functionality has been expanded to cater to the specific challenges faced by English language learners.

We installed RTM in 100 government schools across 6 states and enabled ~20,000 students to use the application. This was done in collaboration with the American India Foundation [AIF], which is a large diaspora led philanthropy focused on accelerating social and economic change in India [www.aif.org]. AIF is one of EH's implementation partners in India. The treatment period was the full academic year from July 2013 to March 2014. The learners used RTM on

average of 2-5 class periods per week on a shared basis. Usage varied as a function of the infrastructure available in the school. In all cases, students shared RTM with other students in a common shared computing environment.

Learners were assessed at the beginning of the treatment period using an assessment instrument designed by EH but administered independently by its implementation partner. EH had no control over administration of the assessment or collection of the data. EH's partner administered the assessment and collected the data. Learners were re-assessed at the end of the treatment period using the same assessment instrument. The assessment was conducted in 88 of the 100 schools. After removing invalid data, we had a total of 4,128 valid responses.

The survey for Grade 8 was administered erroneously and was removed from further analysis. We have controlled for learners who have not been administered both the baseline and endline tests. In the analysis below, we have also excluded outliers with extreme results.

We tested for statistical significance in outcomes using the Wilcoxon signed-rank test [Wilcoxon, 1945]. The Wilcoxon signed-rank test is a non-parametric statistical hypothesis test used when comparing repeated measurements on a single sample to assess whether their population mean ranks differ. We estimate the population to be non-normal and thus the Wilcoxon test is a more appropriate measure of significance.
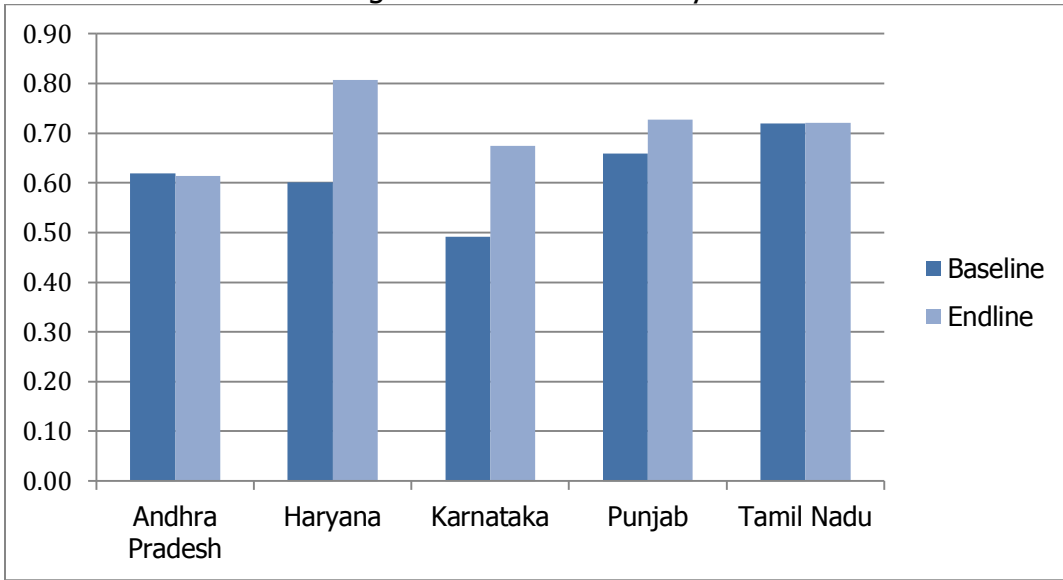
Table 2
Schools by State

State	No of Schools	n[1]
Andhra Pradesh	20	800
Delhi	15	233
Haryana	23	1188
Karnataka	10	737
Punjab	10	441
Tamil Nadu	20	729

[1]n = number of students who took both tests in Classes 6 and 7

Overall Results
Tables 3 and 4 show the average scores by state for Class 6 and 7. In the case of Class 6, a total of 1,580 students in 5 states were administered the same test at the beginning and end of the program. Haryana [+21%], Karnataka [+18%], and Punjab [+7%] showed the greatest gains. Students in Delhi were left out of the analysis because slightly different tests were administered at the baseline and endline.

8

Table 3
Average Scores for Class 6 by State

The frequency distribution of scores for all states for class 6 in Table 4 shows the improvement even more clearly in a significant rightward shift, with endline scores clustered at the top and a mode greater than 90% on the assessment test. The endline scores were stochastically greater than the baseline scores using the Wilcoxon signed-rank test at the 0.05 level.

Table 4
Frequency distribution of Class 6 Scores [All States]
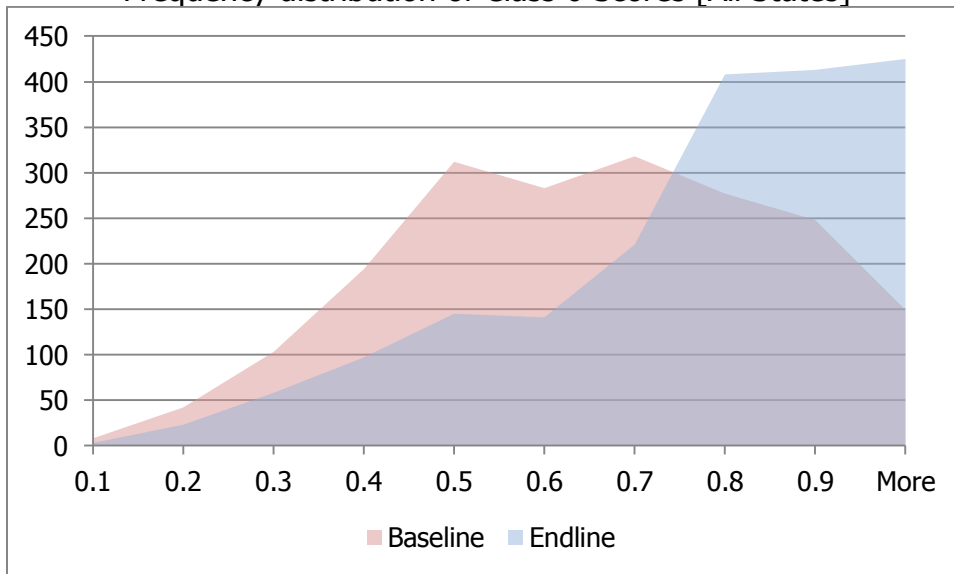
For Class 7, a total of 1,961 students in 6 states were administered the same test at the beginning and end of the program. In contrast with Class 6 results, all 6

states demonstrated significant improvement in their English proficiency at the end of the 9-month treatment period. Haryana [+19%], Andhra Pradesh [+18%], and Punjab [+17%] showed the greatest gains, followed by Karnataka and Delhi.

Table 5
Average Scores for Class 7 by State
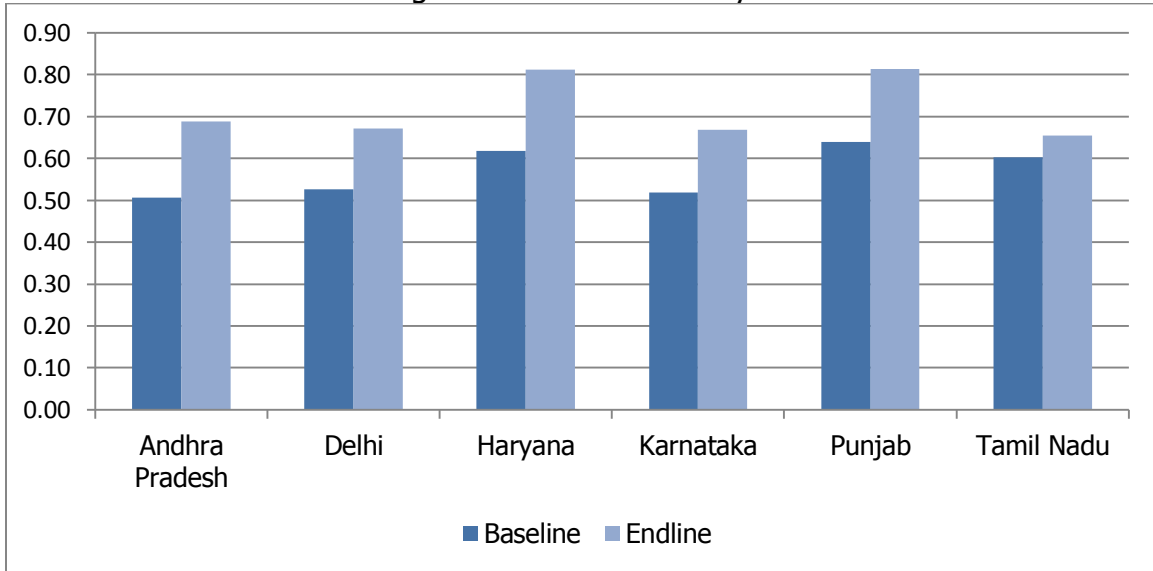
The frequency distribution for Class 7 across states shows a similar marked movement to the right demonstrating that the endline scores were significantly improved compared to the baseline. The difference between endline and baseline scores was significant at the 0.05 level.

Table 6
Frequency distribution of Class 7 Scores [All States]
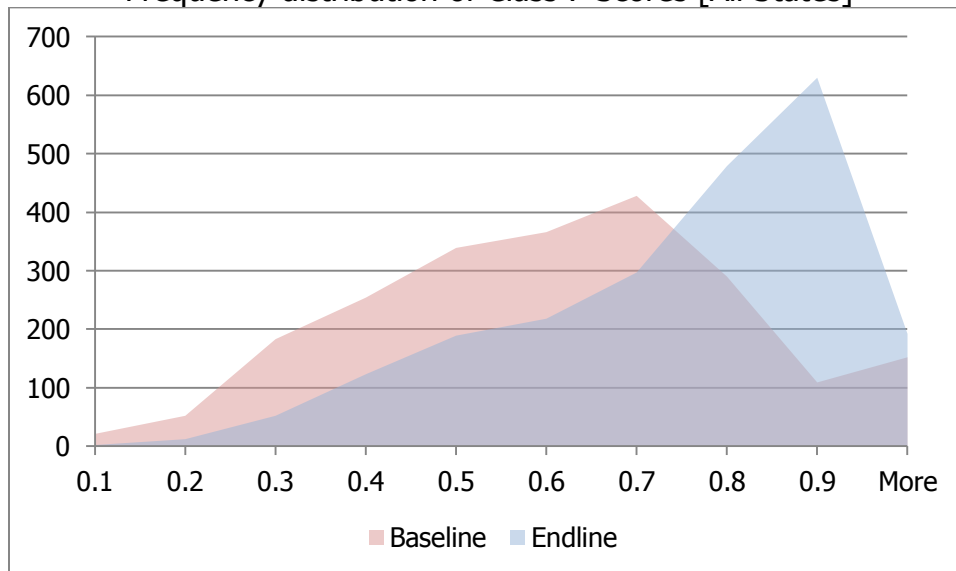
Appendix I lists distribution of outcomes by state for Class 6 and 7. Table 7 below displays the results of the Wilcoxon signed-rank test that the two population mean ranks do not differ. Results for Class 6 in Delhi were excluded because the initial assessment test had half as many questions than the final test. Except in one case, Class 6 in Andhra Pradesh, the endline scores in all the other cases were significantly better when compared to the baseline scores.

Table 7
Statistical significance tests by state

State	Class 6	Class 7
Andhra Pradesh	X	X
Delhi		X
Haryana	X	X
Karnataka	X	X
Punjab	X	X
Tamil Nadu	X	X

We also examined the results by different sections of the survey, which addressed different facets of English language proficiency – reading, reading comprehension, vocabulary, grammar and spelling. These results further confirm that the improvement evident above is also present irrespective of the area of proficiency.

The categories with the greatest number of categories on the assessment were reading comprehension (10 questions) and grammar (8 questions). The endline tests for Class 6 and Class 8 in all states as well as Class 7 in Delhi had different distributions of questions and were therefore left out of the analysis.

Table 8 shows the average scores of the Class 7 students at the baseline and endline for reading comprehension and grammar in Andhra Pradesh, Haryana, Karnataka, Punjab and Tamil Nadu. The difference between the baseline and endline scores in each category was statistically significant using the Wilcoxon signed-rank test at the 0.05 level.

Table 8
Statistical significance tests for Class 7 by category

	H_0	H_1
Reading comprehension		X
Grammar		X

The frequency distributions of both reading comprehension scores and grammar scores display a rightward shift and clustering around the higher range. The grammar scores particularly show a dramatic increase in the number of students attaining high scores.

Table 10
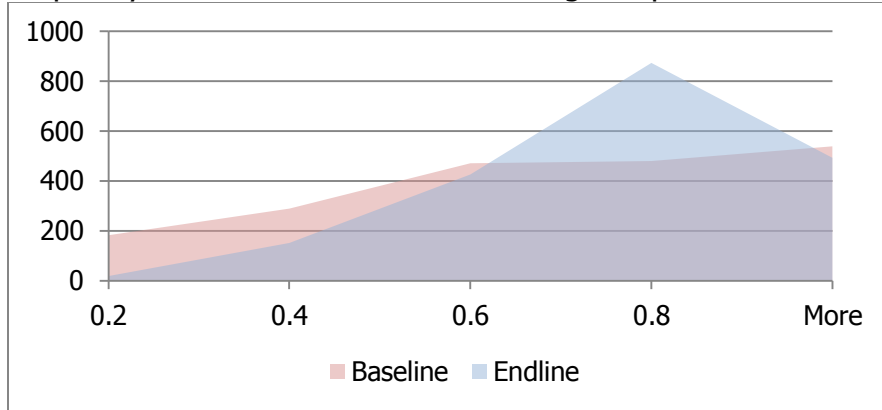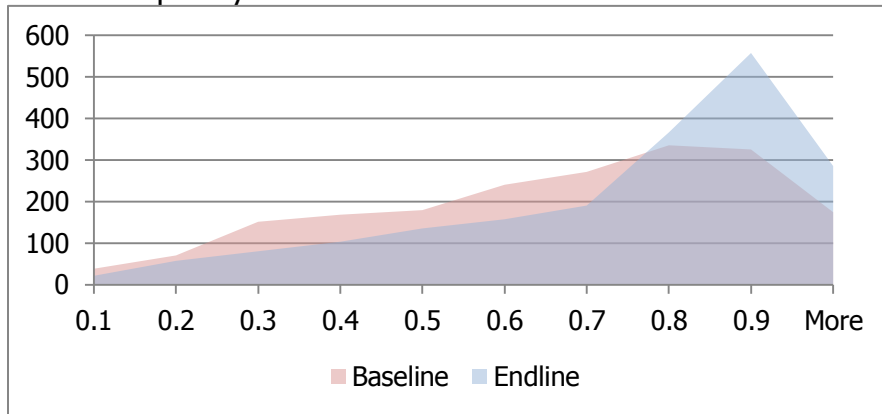Frequency distribution of Class 7 Reading comprehension scores

Table 11
Frequency distribution of Class 7 Grammar scores

Conclusion

This study has provided, perhaps for the first time, large-scale results on the efficacy of multi-sensory technology for language learning and for acquisition of reading skills specifically. We had 20,000+ learners use an English language learning application called ReadToMe™ in a geographically dispersed setting.

The learners came from poor economic backgrounds and the school settings were infrastructurally weak with lack of access to the Internet, shared computers, limited access to the software, and weak classroom facilities. In spite of such limitations, assessments of changes in proficiency levels reveal significant gains across the board in the 6 states at all grade levels and across different facets of language learning – reading, reading comprehension, vocabulary, spelling and grammar. The study provides strong validation that multi-sensory technology can improve English language proficiency levels and reading skills dramatically, with less effort, and potentially in a shorter time period.

We are now expanding the study to a much larger population. We will also be setting up control groups in the next version of the study, which will allow us to assess the impact of the EnglishHelper™ tool in a more controlled manner. Assessments will be performed by independent agencies removing any remaining source of bias in our results.

References

Blomert, L., & Froyen, D. Multi-sensory learning and learning to read. *International Journal of Psychophysiology*, 77, 195-204.

Byrne, B., R, F.-B., 1989. Phonemic awareness and letter knowledge in the child's acquisitions of the alphabetic principle. Journal of Educational Psychology, 80, 313-321.

Cavanaugh, C.S., (2001). The effectiveness of interactive distance education technologies in k-12 learning: A meta-analysis. *International Journal of Educational Telecommunications*, 7(1), 73-88.

Chapelle, C.A. (1997). CALL in the year 2000: Still in search of research paradigms. *Language Learning & Technology,* 1(1), 19-43.

Chapelle, C.A. (2001). Computer applications in second language acquisition: Foundations for teaching, testing and research. Cambridge: Cambridge University Press.

Ghazanfar, A.A., and Schroeder, C.E., 2006. Is the neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10, 278-285.

Hedge, L.V., & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando: Academic Press.

Iandoli, J.J., (1990). CALL and the profession: The current state. *French Review*, 64(2), 261-272.

Liu, M., Moore, Z., Graham, L. & Lee, S., (2001). A Look at the Research on Computer-Based Technology Use in Second Language Learning: Review of Literature from 1990-2000. *Journal of Research on Technology in Education*.

Lopez, O.S., (2010). The Digital Learning Classroom: Improving English Language Learners' academic success in mathematics and reading using interactive whiteboard technology. Computers and Education, 54, 901-915.

Lou, Y., Abrami, P.C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, 71(3), 449-521.

McCandliss, B.D., Cohen, L., Dehaene, S., 2003. The Visual Word Form Area: expertise for reading in the fusiform gyrus. Trends in Cognitive Sciences, 7(7), 293-299.

McCandliss, B.D., Posner, M.I., and Givon, T., 1997. Brain Plasticity in Learning Visual Words. Cognitive Psychology, 33, 88-110.

Mousavi, S.Y., Low, R., and Sweller, J., (1995). Reducing Cognitive Load by Mixing Auditory and Visual Presentation Modes. *Journal of Educational Psychology*, 87(2), 317-334.

Pennington, M.C. (1996). The Power of the computer in language education. In M.C. Pennington (Ed.), The Power of CALL (pp 1-14). Houston, TX: Athelstan.

Pusack, J.P., & Otto, S.K., (1990). Applying Instructional Technologies. *Foreign Language Annals*. 23(5), 409-417.

Salaberry, M.R. (2001). The use of technology for second language learning and teaching: A retrospective. *The Modern Language Journal*, 85 (1), 39-56.

Wilcoxon, F., (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6), 80-83.

Appendix I

Distributions of Scores by State

Andhra Pradesh Class 6

Haryana Class 6

Karnataka Class 6

Punjab Class 6

Tamil Nadu Class 6

Tamil Nadu Class 7

Baseline ■ Endline

Karnataka Class 7

Baseline ■ Endline

Delhi Class 7

Baseline ■ Endline

Punjab Class 7

Baseline ■ Endline

Andhra Pradesh Class 7

Baseline ■ Endline

Haryana Class 7

Baseline ■ Endline