Visual-Syntactic Text Formatting: Theoretical Basis and Empirical Evidence for Impact on Human Reading

Randall C. Walker, MD
Walker Reading Tech., Inc.
Bloomington, Minnesota
rwalker@liveink.com

Phil Schloss, BSEE
Walker Reading Tech., Inc.
Bloomington, Minnesota
phil@red-castle.com

Charles A. Vogel, PhD
Eagle Valley School District
Eagle, Colorado
cvogel@eagleschools.net

Adam S. Gordon
Walker Reading Tech., Inc.
Bloomington, Minnesota
adam.gordon@liveink.com

Charles R. Fletcher, PhD
University of Minnesota
Minneapolis, Minnesota
fletc002@umn.edu

Stan Walker, MD
Northwest Eye Clinic
Minneapolis, Minnesota
dr.walker@nweyeclinic.com

Abstract

Visual-Syntactic Text Formatting (VSTF) algorithms first analyze, then reformat a sentence into cascading patterns that cue syntactic structure and assist visual-processing. VSTF was evaluated in yearlong, classroom-based, randomized controlled trials, with in-class reading sessions (25 minutes per session, twice a week), using electronic textbooks for high school students. Pretest-posttest analysis showed that, in each grade, VSTF students had significantly higher scores on nationally standardized (and conventionally formatted) reading proficiency tests over controls; effect sizes ranged from .41 to .69 standard deviations, and the one-year growth in reading proficiency with VSTF was equivalent to 2 to 3 years' of additional growth in study and national controls. VSTF groups also significantly increased scores, with medium effect sizes, on standardized quizzes and examinations for comprehension and retention of the material in the textbooks. Key words: Text formatting; Syntactic Processing; Reading Comprehensio.

Introduction

The global information economy has promoted both the geographical distribution and the human reading of text in electronic, rather than in print form. Individual customization is also an important new advantage of digital text display; for example, many readers already take extra steps to increase the size of the font, or narrow the column width of their electronic texts to improve their reading experience and performance.

Visual-syntactic text formatting (VSTF) is a new text formatting method that can be integrated with digital text systems and publications as an alternative format that individual readers may find helpful.

This paper will review: the major principles of visual and syntactic processing upon which the VSTF method is based: key findings from recent controlled trials that evaluated the impact of VSTF on reading performance: and some examples of how VSTF is now being deployed and integrated with commercial text and publishing systems.

.

Theoretical Background

Representation of Syntax Structure in Speech and its Transcripts. A wide range of neurocognitive, linguistic, and psychological research affirms that an important dimension for the representation of meaning in natural spoken language is syntax. However, syntax is more complex than a simple, concatenated sequence of one phrase after another; rather, it is hierarchical, much like a set of Russian dolls, in which smaller dolls, or phrase-groups, are "nested" inside ever-larger ones. The human mind's capacity to build sentences through the recursive process of nesting language units inside other units, and thereby transforming them, is the essential feature that enables human language to represent an infinite number of meanings [1].

When natural language is spoken, it is produced, perceived, and interpreted as a linear structure -- through time -- which limits its capacity for conveying the multi-dimensional, hierarchical structures of syntax. Nevertheless, this linear structure can be enriched with prosodic cues, which in turn can denote syntactic relationships to enhance the efficiency of a listener's comprehension of a spoken sentence. Prosodic cues,

1-4244-1243-9/07/$25.00 ©2007 IEEE

which give speech a highly differentiated acoustic structure beyond the acoustic representations of the words themselves, are more subtle and multidimensional than the simple pauses that occur at major phrase boundaries, and have been shown to be powerful enough to enable listeners to accurately predict the syntactic categories of about-to-be uttered phrases, based on prosodic patterns leading up to the not-yet-uttered phrase [2] The interaction between prosodic cueing and syntactic processing during listening has been shown to dynamically affect the processes needed to attain efficient sentence comprehension [3]. Moreover, when prosodic structure is experimentally stripped away from digital audio recordings of speech, listeners' comprehension drops [4].

Up to the present time, the transcription of natural language sentences has also been linear. Within such linear scripts, some specific cues, (such as punctuation marks), denote syntactic boundaries; some (but not all) of these punctuation marks correspond to pauses and prosodic variations in spoken language.

However, when sentences become longer and more complex, working memory is overloaded, and the efficiency of comprehension can break down. (For example: *The parent of the student who got lost during the class field trip to the capitol called the principal*.) This "memory bottleneck" in language comprehension appears to arise because "hierarchical linguistic relations must be recovered from a linear input stream" [5].

Visual Processing while Reading Block Text . Given the conventions of writing natural language in a one-dimensional linear sequence, as a direct "transcript" of spoken language that is uttered through one-dimensional time, it is not surprising that conventions for reading research have also been predicated on the visual perception of words aligned in linear sequences, across multiple rows.

As the first steps in reading, the visual perception and linguistic interpretation of words written in linear sequences, across multiple rows, are neurocognitive processes that interact in unique ways compared to the perception and interpretation of natural scenes [6, 7, 8]. In reading text in blocks of rows of words, attentional processes are strained as fixations move from one word to the next, trying to ignore those words that are above and below the targeted word, but which compete for attention while having no logical relationship to the targeted word (See Figure 1) [9, 10]. Interpretation of the next word in a sequence often encounters a "garden path" effect that requires regression to previously viewed words [11].

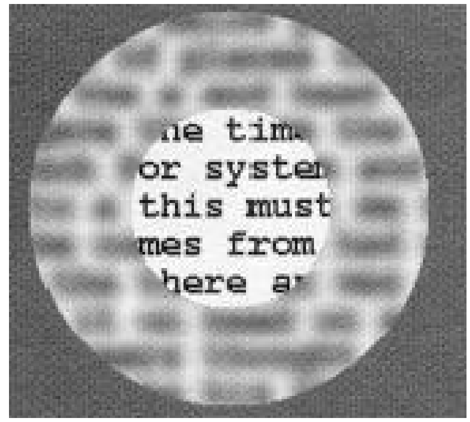


Figure 1. Relative Visual Resolution at a Typical Fixation Point When Reading Block Text. ***A suurounding area of lower-resolution visual data can be perceived around the higher-resolution, central (foveal) area.***

Even the size of the parafoveal zone, (which otherwise has lower resolution visual data that can help direct the landing spot for the next fixation), shrinks when the difficulty of a viewed word, or the complexity of syntactic relationships that a viewed word might have, increases -- thus impeding the efficiency of reading all the more [6].

However, in contrast to the linear-temporal medium of spoken language, visible language media, including the formatting of words on a two-dimensional display surface, are able to represent more dimensions of information than rows of words, and are also able, at the same time, to follow conventions (e.g., left to right, top to bottom) that still unambiguously maintain the linear sequence of the original natural language sentence. Moreover, the dynamic perception of such "extra-linear" structure that is built across the display's additional dimensions is able to harness more natural, and more powerful mechanisms for perceiving natural scenes and patterns [12, 13, 14].

There is evidence that users of text already use such "extra-linear" structural information to improve their comprehension of a document: they examine how many pages it has, look at figures and chapter titles, and check the index -- all before reading any of the text in a serial, linear fashion. Moreover, even within passages, it is known that first examining diagrams and tree structures of the concepts covered in the text itself are activities that increase reading comprehension of the text [15, 16, 17].

Visual-Syntactic Text Formatting. The purpose of VSTF is to make a sentence diagram that can similarly increase reading comprehension of the sentence itself; but it does not simply make a diagram *of* a sentence -- rather, it makes a diagram *with (and only with)* the words of a sentence, positioning the words and segments into specific locations relative to one another to create "extra-linear" structural information.

When in the Course of human events, it becomes necessary for one people to dissolve the political bands which have connected them with another, and to assume among the powers of the earth, the separate and equal station to which the Laws of Nature and of Nature's God entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the separation.

Figure 2a. The First Sentence of the Declaration of Independence Presented with a Conventional Block Text Formatting Method.

This extra-linear spatial structure can be initially examined, before the specific, word-by-word linear interpretation of the text. In addition, the visual contours of multi-segment cascades are designed to provide low-resolution visual cues that can be dynamically perceived in the parafoveal zone, even as the targeted word at each fixation is being directly examined and interpreted. These low-resolution parafoveal cues can denote the hierarchical syntactic relationships that exist between adjacent phrase-segments on separate rows, and even across remotely separated rows. These cues, in turn, can interact with the other dimensions of information-flow that are each dynamically combining and collaborating to enhance sentence comprehension [18, 19].

In this way, reading, unlike listening, can harness the multidimensionality of its visual medium, to dynamically represent and integrate processing of hierarchical linguistic relations that would otherwise, in speech (or in block text), need to be recovered from a linear input stream; in this way, it is possible to open up a "memory bottleneck" in language comprehension that otherwise constrains the efficiency of both listening and reading linear block text.

Moreover, by providing such "extra-linear" structural information in a transparent manner, without additional lines, symbols, or characters, this formatting method can be read continuously, without distracting or fatiguing the reader. Thus, student readers can also be exposed to a wide variety of sentence structures, each depicted in an explicit, but transparent, visual form, that enhance the student's syntactic processing and awareness, and thereby, with continued practice (such as reading one's expository reading assignments in a core content area class), strengthen long-term reading proficiency [20, 21, 22].

Empirical Evidence for Impact of VSTF on Reading Performance

The VSTF method has already been studied in laboratory settings, using within-subject designs in adult readers, and in a yearlong, classroom-based controlled experiment,

When in the Course
of human events,
it becomes necessary
for one people
to dissolve the political bands
which have
connected them with another,
and to assume
among the powers
of the earth,
the separate and equal station
to which
the Laws of Nature
and of Nature's God
entitle them,
a decent respect
to the opinions
of mankind
requires
that they
should declare the causes
which impel them
to the separation.

Figure 2b. The First Sentence of the Declaration of Independence Presented with the Visual-Syntactic Text Formatting Method.

involving ninth grade high school students [23]. This prior research demonstrated that VSTF can increase reading performance (increasing immediate and long-term retention of the material read in the VSTF), while also strengthening students' performance on nationally standardized (and conventionally formatted) reading comprehension tests.

In this paper, we will report additional school-based studies, which corroborated and expanded upon these initial observations.

Yearlong, Randomized Controlled Trials in Classroom Settings

METHODS: Research Site. The site for this research was a high school in the western United States. Because this research involved the evaluation of educational interventions in a classroom setting, the research qualified for an exemption (category 1) from Federal Human Subjects research regulations. All student data were kept within the school district, and were analyzed exclusively by school district personnel. The research spanned the 2003-2004 academic year.

Student Demographics. The proportion of students for whom English was a non-native language was approximately 35% in each grade. The percentage of students qualifying for free- or reduced- price lunch programs was 27.5%. Students who were taking Advanced Placement or Community College classes in History for college credit, and thus not taking the standard History curriculum, were studied separately; these represented approximately one-fifth of the students in each grade. Demographic data and reading proficiency test results on VSTF treatment and control students were collected prospectively, including district and state-mandated nationally standardized reading tests that were administered to students.

Teacher Participants. All three of the high school's social sciences teachers participated, each teaching both a control and an VSTF treatment group in the same grade. All participating teachers received basic orientation and training on the design and implementation of the study; however, there was otherwise no professional development for teachers on methods to improve reading performance with their students. One of the participating teachers provided on-site management of the study, and was the district contact for student data collection and analysis, keeping all raw data within district property**.**

Materials *Texts and Reading Sessions.* The main texts were the History/Social Sciences textbooks for each grade (10 and 11). Both VSTF and block text versions of electronic textbooks were prepared on software platforms using the Microsoft Windows® operating system. Both VSTF and Block Text electronic textbook platforms permitted font enlargement, choice of dark or light background colors, and point-&-click table of contents. Block formatting for the control groups' electronic textbooks used the same number of characters per line as was found in the standard, paper-based textbook, with left margin-only justification. (See Figure 3, VSTF and Block formatted electronic textbooks).

Computers. Laptop computers using the Microsoft Windows® operating system were used, and the electronic textbooks were presented in Windows-based applications. Laptops were kept on a rolling cart that could be moved from classroom to classroom. Reading sessions could also be conducted in the school's computer labs, which used desktop PCs.

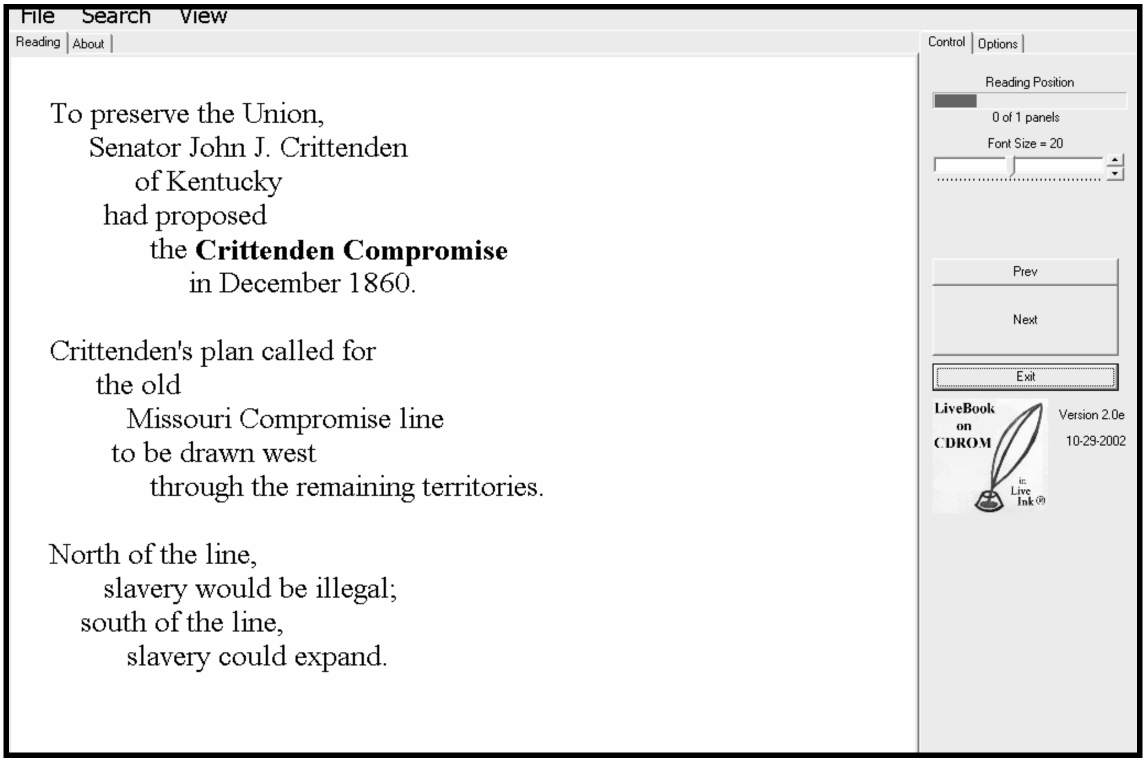
Study Management and Support. One of the participating teachers provided on-site management of the study and was the district contact for student data collection and analysis, keeping all raw data within district property and providing the study sponsor with the final data analyses reported in this paper. The study sponsor, the technology provider for the electronic textbooks and VSTF formatting, was located over 1000 miles from the study site, and did not participate in the management of the study, making only one site visit before the study began to orient teachers to the technology. The school district's Information Technology staff provided on-going technical support for computer labs, laptops, Internet connections, and software installation; these IT staff would contact the study sponsor as needed for technical support.

Study Design: *Randomized Controlled Trials (RCT).* Students taking standard history courses in 10th grade (n=84) and 11th grade (n=60) were randomly assigned, using the school district's class scheduling software, to one of six (in grade 10) or four (grade 11) different class sections of History. The district used the scheduling software explicitly to achieve such randomization effects, in order to prevent "clustering" of student subgroups (e.g., groups of friends) to a particular class section or teacher. Except for the exclusion of the students who were taking the Advanced Placement and Community College History courses, no other variables were included in the software program to assign students to class sections or teachers.

There were 3 teachers for grade 10, each having two grade 10 sections. Each teacher was to have one of their sections assigned to the VSTF treatment and the other section assigned to the control condition; this assignment was made by coin-toss randomization by a district administrator not affiliated with the study, after individual student-to-class section randomization had been made. Similarly, for grade 11, there were 2 teachers (who also taught grade 10 classes). Each of these grade 11 teachers also had one VSTF treatment and one control class-section; randomization by coin-toss performed by a district administrator not affiliated with the study assigned one of the teacher's two class sections to the VSTF treatment condition, and the teacher's other section in that grade to the control condition. Rarely, section transfers were requested by individual students (before randomization of a section to a treatment condition) for personal scheduling reasons, but never exceeded one student per section.

Importantly, all of the teachers, in both grades 10 and 11, for both their individual VSTF classes and control classes, coordinated their curriculum content coverage and student testing in complete uniformity, (both between their own two types of classes and among each other), covering, week by week, identical chapters and sections for the course textbooks, which had been agreed upon in advance. All testing instruments, including quizzes given after each reading session, unit exams, and final exam were identical across all teachers at each grade, and for both VSTF treatment and control class sections.

Intervention Method. The intervention method consisted simply of having students read their assignments for their History course in a classroom setting. At each grade, students read during class time for approximately 50 minutes a week, usually 25 minutes within a class session that would last overall for 90 minutes every other day. For both VSTF treatment and control groups, the teacher-supervised reading sessions occurred in History classes. No electronic textbook reading occurred outside of class time; students were still free to use their paper textbooks for homework assignments. Because neither the VSTF electronic textbook nor the block electronic textbook had any images or figures, students in both the VSTF



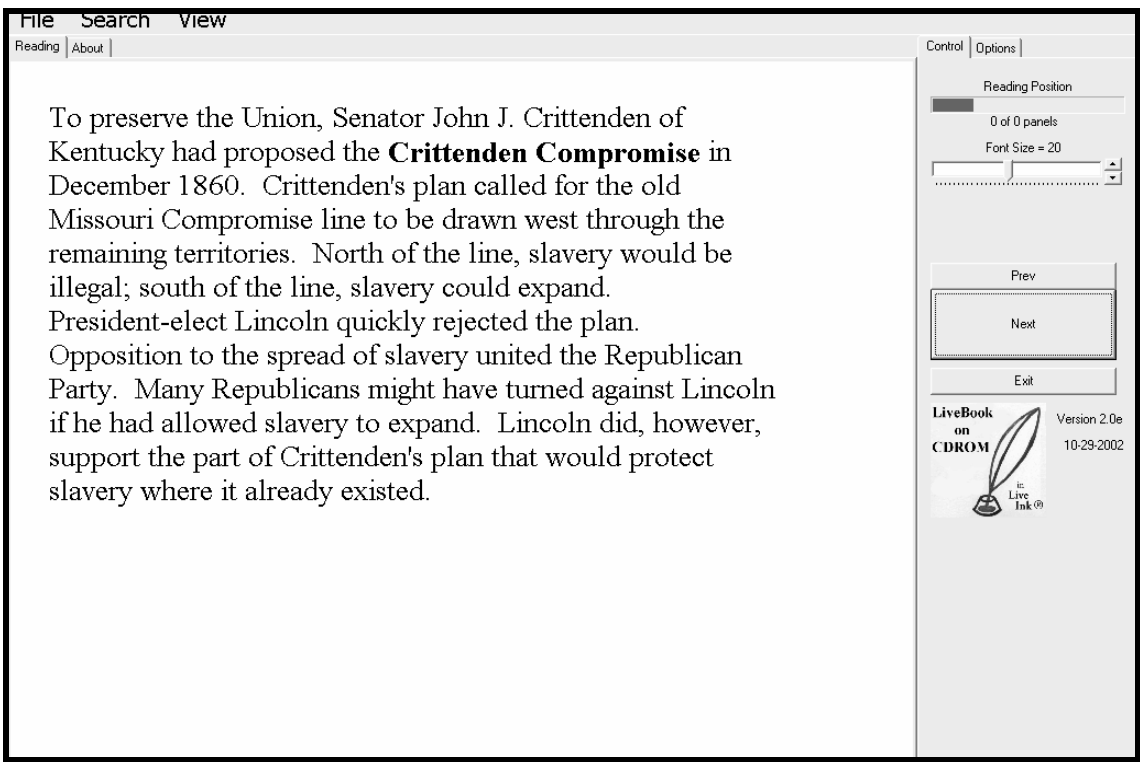


Figure 3 . Electronic Textbooks Used in VSTF Treatment Group (Above) and Control Group (Below). ***Both platforms offered optional: dark background with light text, font enlargement, search, and point -&-click table of contents. The VSTF book required about twice as many clicks to advance through a passage as the Control.***

treatment group and the control group were also always free to open and use their standard, paper-based textbooks in class, either to complete the assigned reading section or to examine the paper textbook's figures and images.

After each History reading session, students were given a 10-point quiz, (obtained from the publisher-provided test-generator for the textbook), which counted toward a student's grade. Because participation in the study was integrated with classroom activities, and thus controlled directly by teachers, there was no attrition, and no crossover between study groups. Fidelity to the study was validated with teacher-logs, and from the records of the quizzes themselves, which were given to students after each reading session.

Outcomes Measurement: ***Academic Achievement*** For both grades 10 and 11, it was also possible to analyze the scores of students' quizzes given after each reading session, and of students' unit exams, given approximately every 3 weeks during the year. For grade 10, (World History), there were 38 quizzes and 10 unit exams. For grade 11, (US History), there were 48 quizzes and 16 unit exams. All of these quizzes and examinations were uniform across all teachers and conditions. The quizzes and exams were generated by test-generator software that publishers routinely provide with their textbooks.

Outcomes Measurement: ***Transference to Reading Proficiency in Standard Format.*** For several years prior to the study, the school district had already adopted the Northwest Education Association (NWEA) Measure of Academic Progress test for Reading, as a tool to assess individual student progress and to document teacher performance [24]. For this study, for grade 10, both in the control groups and VSTF treatment groups, the school district's data on the NWEA test was used for both pretest, (baseline), measurement of reading comprehension in the early fall, and for end-of-year, posttest, measurement of reading comprehension in the late spring. NWEA testing was not available for 11th grade students. Therefore, for pretest measurement of 11th grade students, district-based data of each student's score on the state's Survey of Academic Progress (SSAP) reading section was analyzed, using scores from the previous year; (i.e., scores from SSAP tests given in the winter of 2003, when student's were in grade 10, were used as pretests for students who were in the 11th grade during the 2003-2004 academic year). Because the district required all students to take the ACT college placement test in the spring of their 11th grade, (whether the student planned to go to college or not), it was possible to use the students' ACT reading section scores as a posttest measure of 11th Grade reading comprehension.

As an additional, categorical independent variable for 10th grade analysis, seven levels of district-assigned reading proficiency were also used; these proficiency levels were based on district-administered tests given in the winter of 2003, when students had been in grade 9. Importantly, all reading comprehension test passages, (NWEA, SSAP, ACT), were formatted in a conventional block formatting pattern, with the tests being taken directly as provided by the testing publishers.

Statistical Methods: ***Baseline Equivalency***. The equivalency of VSTF treatment and control groups, at each grade, was determined by performing a one-way, between-groups ANOVA. For grade 10, equivalency testing used the Fall NWEA test. For grade 11, equivalency was determined using the SSAP results from of the winter of the previous school year.

Reading Comprehension. The null hypothesis in these experiments was that use of the VSTF treatment would have no effect on nationally-standardized Reading Comprehension scores. Pretest-posttest analysis of all grades was performed with an analysis of covariance, (ANCOVA), using post-test Reading Comprehension scores (NWEA for grades 10, ACT for grade 11) as the Dependent Variable, and pretest Reading Comprehension scores as the Covariate, Independent Variable.

To further adjust for the possible effects, (despite the randomization and matching procedures), of undetected clustering of factors that, in turn, might have affected posttest scores and/or response to the VSTF treatment, several categorical pretest variables, (reading proficiency level, non-native English language status, gender, high school teacher or middle school team, and various combinations of these factors), were included as Independent Variables in the ANCOVA.

The same ANCOVA with multiple categorical pretest Independent Variables was also used for analysis of individual *growth* on the NWEA test (i.e., each student's NWEA Fall to Spring difference), for grade 10.

Group Comparisons on Quizzes and Exams. The null hypothesis in these experiments in Grade 10 and 11 was that the use of the VSTF treatment would have no effect on cumulative academic achievement, as measured by all quizzes and unit exams given over the academic year. To test this hypothesis, scores of Quizzes and Exams given throughout the academic year in the 10th and 11th grade, were analyzed with a Test of Repeated Measures multivariate analysis of variance, with pretest reading comprehension as an independent variable.

Assessment of Practical Educational Impact. Several calculations were used to represent the practical significance of changes observed, including standardized effect size differences (Cohen's d, using pooled SD from control and VSTF treatment groups) [25], differences in nationally standardized percentile rankings, and growth in reading comprehension relative to average annual growth in nationally standardized groups.

Statistics Software. Statistical computations were performed using the Statistics Package for Social Sciences, version 10.0.

RESULTS: *Pretest Equivalency.* The studies for grades 10 and 11 were randomized controlled trials; although there was a possibility of pretest differences between VSTF treatment and control groups, as a function of chance, the two groups were nonetheless found to be equivalent at baseline, and had similar proportions of gender and native language subgroups in each group.

Results on Reading Comprehension Tests. Across both grades studied, the VSTF treatment condition was associated with statistically significant, and academically meaningful, pretest-posttest improvements in reading proficiency over controls, with medium to large effect sizes. (See Table 1).

As stated in the Methods section, several categorical variables were included as Independent Variables in the ANCOVA, in addition to the quantitative pretest Reading Comprehension score. Even with these multivariate adjustments, the ANCOVA demonstrated highly significant differences, all with $p < 0.005$ (Tables 2 and 3). The practical and educational significance of the increases is also evident when national percentile ranking scores are used (See Table 4): whereas all groups were equivalent on the pretest measure, the median posttest scores for VSTF treatment groups were 16 to 18 national percentile-ranking points higher than controls. In grade 10, where the NWEA was available for both pretest and posttest, individual student growth in NWEA was also significantly higher in the VSTF treatment group compared to controls. Whereas growth in the control group was on par with an expected, 1-year's worth of growth compared to the NWEA's national standard controls, growth in the VSTF treatment group was equivalent to 2 to 3 years' worth of such national control group growth.

In Grade 11, all students in the district were required to take the ACT, a college entrance exam, even though many did not plan to attend college, and the study participants did not include 17 students taking a Community College History course for college credit. Therefore, ACT performance in study participants would likely be below the national average, which is based on use as a college admissions requirement. Despite the equivalency of the 2 groups at pretest, the pretest-posttest ANCOVA showed a significant increase in ACT scores among the VSTF students, whose ACT national percentile ranking of 51.9 actually surpassed the national 50^{th} percentile (of college-bound students), and was 13 percentile points higher than the control group.

Results on Quizzes and Unit Exams. The RCT design made it possible to compare scores on quizzes and exams that were given regularly, and simultaneously, in both conditions throughout the academic year. The VSTF treatment groups had significant increases for both types of tests, with medium effect sizes (Tables 1). A graph illustrating the transformed variable averages of the VSTF treatment and control groups, for unit exams, in grades 10 are shown in Figure 5.

<u>*Non-native English Language students.*</u> Non-native English language students (ESOL) using the VSTF treatment had significantly higher scores than ESOL students in the control groups. Most ESOL students were recent (less than 3 years) immigrants from Mexico, although some had had English as a second language since early childhood.

Table 1. Mean Scores, SD, and Effect Sizes of Differences Between Control Group and VSTF Treatment Group on Pretest-Posttest Reading Proficiency Tests, Quizzes, and Unit and Final Exams

Grade		**Block Text Control**		**VSTF Treatment**		**Student Level Effect Size**	**Percentage of VSTF Group with Scores above the**
(n)	**Measure**	**Mean**	**SD**	**Mean**	**SD**	***Cohen's d***	**Control Mean**
10	*Pretest NWEA*	226.09	(9.9)	227.28	(11.8)	0.11	54
(84)	*Posttest NWEA*	225.8	(9.9)	233.23	(11.1)	0.69	74
	Growth NWEA	-0.295	(6.48)	6.282	(7.189)	.96	82
	All 38 Quizzes	64.31	(20.9)	72.02	(19.7)	0.38	65
	All 10 Unit Exams	68.1	(20.1)	76.56	(16.26)	0.47	70
	Final Exam	63.93	(15.88)	81.18	(12.32)	1.22	92
11	*Pretest SSAP10*	679.62	(35.33)	690.94	(44.25)	0.28	59
(60)	*Posttest ACT*	18.35	(5.01)	21.06	(4.46)	0.57	73
	All 48 Quizzes	76.38	(17.62)	81.65	(15.46)	0.32	64
	All 16 Unit Exams	53.73	(16.41)	61.18	(15.88)	0.46	68
	Final Exam	83.48	(12.12)	90.03	(8.75)	0.63	78

Grade 10 Dependent Variable: NWEA-Spring Source:	Type III SS	df	Mean Square	F	Sig.
Corrected Model R Squared/Adjusted	9354.620 .921/.802	50	187.092	7.707	.000
Intercept	347.892	1	347.892	14.332	.001
Pretest NWEA	133.872	1	133.872	5.515	.025
Error	801.046	33	24.274		
Total	4421614.000	84			
Corrected Total	10155.667	83			

Table 2. ANCOVA: Between-Subjects Effects, Randomized Controlled Trial, Grade 10
Quantitative Independent Variable (Covariate): Pretest NWEA Reading Comprehension. Qualitative (non-metric) Independent Variables: Gender, Native Language Status, Teacher, Pretest State Reading Proficiency Level.

Grade 11 Dependent Variable: ACT Source:	Type III SS	df	Mean Square	F	Sig.
Corrected Model R Squared/ Adjusted	747.695 .530/.339	17	43.982	2.784	.004
Intercept	67.800	1	67.800	4.291	.044
Pretest SSAP-10	258.270	1	258.270	16.347	.000
Error	663.555	42	15.799		
Total	24815.000	60			
Corrected Total	1411.250	59			

Table 3. ANCOVA: Between-Subjects Effects, Randomized Control Trial, Grade 11.
Quantitative Independent Variable (Covariate): Pretest State-Standards Test for Academic Progress (SSAP) Reading. Qualitative (non-metric) Independent Variables: Gender, Native Language Status, School, Teacher.

These students, as a group, had significantly lower pretest reading comprehension scores than native English-speaking students, but had relative gains with the VSTF treatment that were comparable to the gains seen in the native English-speaking students using the VSTF treatment. In effect, the gains in reading comprehension with VSTF closed about half the gap seen between native and ESOL students at baseline (see Figure 4).

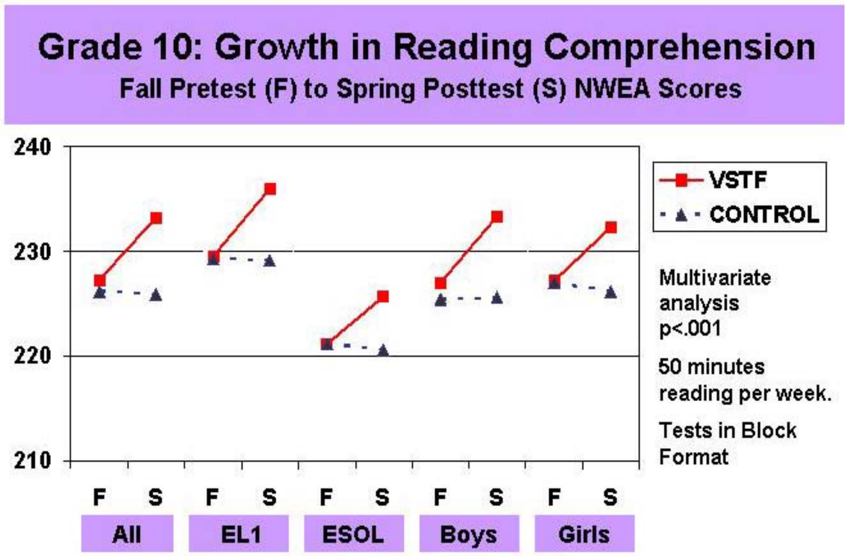


Figure 4. Pretest-Posttest Scores, from Fall (F) to Spring (S) on the Reading Section of Measure of Academic Progress Test by Northwest Educational Association (NWEA). Comparison of Pretest-Posttest Scores in Control and VSTF Treatment Groups, in Native English (EL1) and Non-Native English (ESOL) Students, and by Gender

In both Grades 10 and 11, ESOL students using VSTF also had proportionate gains in quiz scores and unit exam scores that were comparable to the gains seen among native-English speaking students in VSTF groups.

Impact of Prolonged Use. The impact of the VSTF treatment increased with longer use, particularly among lower proficiency students and ESOL students, in whom quiz scores did not rise above controls until after about 6 to 8 reading sessions. By contrast, students with higher baseline reading proficiency, and/or were native English-speaking, appeared to experience an immediate benefit from VSTF on quizzes and exams. Similarly, among all groups, the difference between VSTF treatment and controls in scores on unit exams during the first half of the year (effect size = .375) was not as great as during the second half of the year (effect size =.55), suggesting again that the impact of the VSTF treatment may increase with longer use.

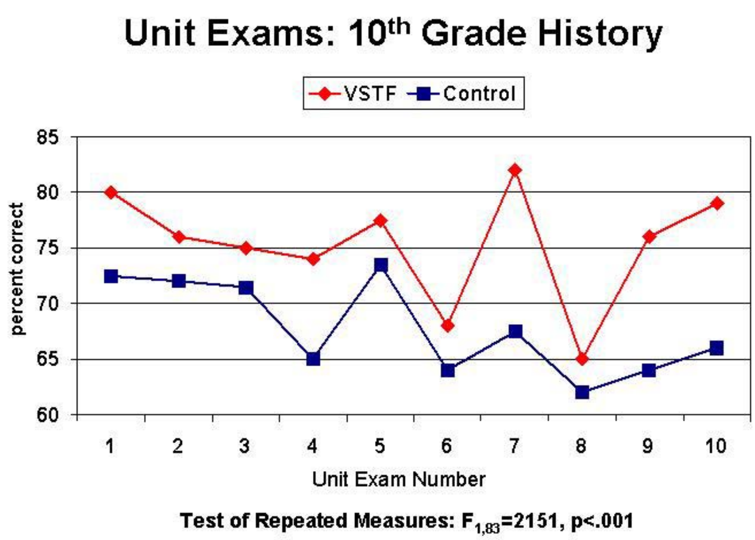


Figure 5. Tranformed Variable Averages for Unit Exams in 10th Grade History Students.
The Unit Exams were given every 3 to 4 weeks during the academic year

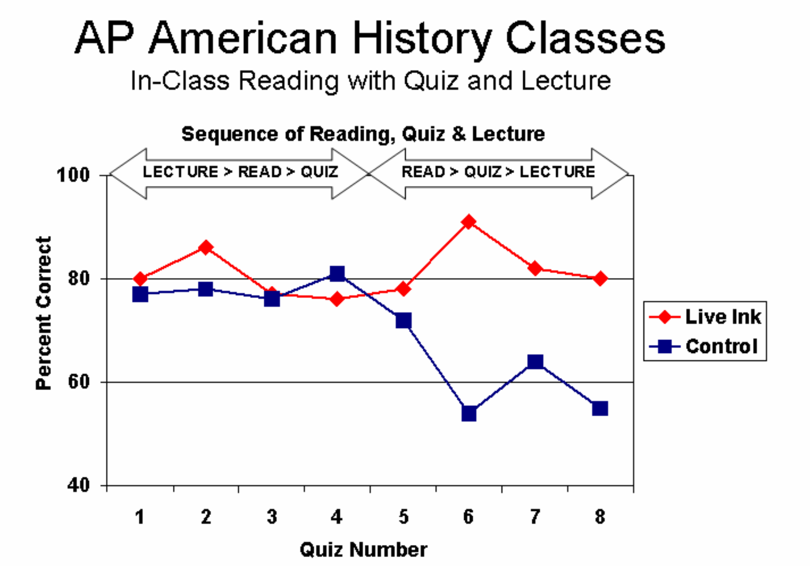


Figure 6. Scores on Quizzes Given to Two Equivalent Groups of Advanced Placement Students, Immediately After Reach Reading Session. ***For Reading Session and Quizzes 1-4, a Lecture was given before both reading and the quiz; for Reading Sessions and Quizzes 5-8, the lecture was given after both the reading and the quiz.***

Pilot Study: Advanced Placement Courses

An Advanced Placement (AP), study evaluated two equivalent groups of American History students; equivalency was determined by a PLACE test that guaranteed that students could read at a college level. The first half of the year, all students read regular textbooks independently, taking quizzes periodically. In the second half of the year, in-class reading began, one of the class sections read in the computer lab with VSTF, while the other class read from the paper textbook. Quizzes (thirty multiple-choice questions per section read) were from a test generator provided by the textbook company, and were administered directly after an in-class reading session. The data points in the graph in Figure 5 represent scores for the second semester only. For first 4 quizzes of the second semester, students from both the control group and the VSTF group were listening and taking notes from a teacher's lecture about the section topic; then the readingwas used as a backup to the lecture and the quiz came right after the reading session (LECTURE-READ-QUIZ). For quizzes 5 through 8, the teacher reversed the process, having students in both groups read and take a quiz first, and then the teacher lectured to the section as a backup (READ-QUIZ-LECTURE). The experiment showed that, for quizzes 5-8, without the auditory, traditional instruction, the AP students in the control group dropped quiz score averages, dramatically. By contrast, the students in the VSTF (Live Ink) group maintained high quiz scores even without having the lecture before the quiz. The graph in Figure 5 was generated from a repeated measures analysis; the ANOVA compares between subjects effects for every data point during the second semester experiment, regardless of the lecturing sequence. (VSTF versus Control for all 8 quizzes: $F= 28.891$, $p<.001$, ANOVA, between subjects).

GRADE POSTTEST	**10 NWEA**		**11 ACT**	
GROUP	**CNTL**	**VSTF**	**CNTL**	**VSTF**
(N)	(40)	(44)	(31)	(29)
Mean	51.95	65.38	38.66	51.94
SEM	3.11	3.93	4.88	3.93
Median	54.00	72.00	31.00	47.00
SD	20.61	24.87	26.28	21.85
Minimum	2.00	9.00	6.00	11.00
Maximum	87.00	98.00	90.00	90.00

Table 4. Posttest Reading Proficiency Scores, as National Percentile Rankings. ***Groups (Control CNTL) and VSTF were equivalent at Pretest testing. These Grade 11 students did not include those taking Advanced Placement History.***

Discussion

These two randomized controlled trials demonstrated that the direct use of the VSTF method in textbooks enables students to score higher on the quizzes and exams that textbook publishers provide to assess academic achievement in content areas covered by the textbooks. At least 6 reading sessions may be required before poorer readers and/or non-native English learners experience a measurable benefit from VSTF, whereas better readers appear to benefit with the first session.

These RCTs also demonstrated that regular, frequent (2 or 3 times per week), but relatively short duration (25 minute) reading sessions, over 8-months in a standard academic year, in which the reading material is reformatted in visual-syntactic cascading patterns, can improve students' reading proficiency, as measured by a nationally standardized test (NWEA) commonly used to assess Adequate Yearly Progress, as well as by college entrance examination (ACT) reading tests. Because such standardized tests are formatted in a conventional block-text format, this regular use of VSTF text appears to transfer to a long-term, format-independent increase in reading proficiency. This is especially notable because, on a national level, year-to-year progress in reading proficiency diminishes to nearly zero by the 9th grade.

This research also found that, compared to native English language students in the control groups, non-native English language students in the VSTF treatment groups closed about half the gap in reading proficiency found between these groups at pretest testing. Both students with higher pretest reading comprehension and with lower pretest reading comprehension scores had significant improvements in reading comprehension with VSTF compared to controls. Boys and girls benefit to a similar degree. Pilot studies in high school students with relatively high proficiency compared to their peers, but who nonetheless can be challenged when reading college textbooks, also appear to benefit from VSTF.

Limitations of the current study.

This study involved only a single, but demographically diverse, rural-suburban school district in the southwestern area of the United States. Although it has several demographic variables that are common in US schools today, additional research is needed to confirm that similar results can be seen in more extreme conditions, such as urban areas with higher rates of school-lunch program students, or a higher rate, or more diverse range, of non-native English students, (i.e., other non-English languages besides Spanish).

The differences seen are unlikely to be the result of a Hawthorne effect from the novelty of reading from computer laptops, because both the VSTF treatment and control groups had access to laptops in the classroom throughout the study. Students in both of these "intent to treat" groups also had the option of reading from their paper textbooks at any time; for example, if they wanted to examine charts or maps. Notably, after several months into the study, large numbers of students in the control group, but none in the VSTF treatment group, began to read from the paper textbooks only, letting their available laptops remain on the storage cart. These control students reported, in a survey given several months into the study, that it was the lack of images in the laptop version of the textbook that led to their preference for the paper textbook. However, the students in the VSTF treatment group, who did not go back to their paper textbooks, (although they always had the option to do so), did not complain of such lack of images in their VSTF laptop textbooks. The differences we observed between the VSTF treatment group, (who used laptop textbooks all year, even without images), and the control group, (who, despite having yearlong access to laptops, preferred paperbooks), are unlikely to be the result of less time on task, because both groups had identical schedules for teacher-supervised, in-class reading sessions. Finally, the differences observed are also unlikely to be the result of control group demoralization, because the control group in the 10th grade nevertheless performed as well as 10th grade national controls on the NWEA.

In this RCT, the high school students in both VSTF treatment and control groups were permitted, if they wished, to read passages aloud, in a low voice, as long as neighboring students did not object to the practice. Many students in the VSTF treatment group, particularly lower level or non-native English students, adopted this practice of reading aloud softly, at least for their initial reading of the text. By contrast, very few of the students in the control groups adopted this practice, and even resisted to do so when this was suggested. It is possible that some of the benefits seen with the VSTF group, particularly in lower level readers, were due to a "visual scaffolding" effect that supported students' reading aloud, a practice that, in turn, was the primary agent for the observed increase in reading comprehension. However, the VST method also improved scores among better readers, who usually did not adopt the reading aloud practice.

Practical Implementation

This research also demonstrates that classroom-based delivery of the VSTF treatment is feasible. The computer systems, (sets of laptops on rolling carts), used for in-class reading are already widely available from major educational technology providers, and their prices continue to decline. The VSTF method will also be easily implemented on yet lower-cost educational "appliances" now in development, on mobile, hand-held computers and video media players, and larger screen cell-phones.

The primary purpose of using in-class reading sessions in this research was to assure that students actually read their texts and took quizzes immediately after their reading sessions. Other than orientation of participating teachers to the logistics and implementation of the study, there was no professional development or teacher training in this study that provided teachers with new methods to improve student reading that were based on the VSTF format. However, even though this study did not examine the effect of building a curriculum around the possibility of using VSTF, it is possible that particular teaching methods , (not only in textbook reading, but for all types of reading and writing that involve digital texts), could be developed that will achieve greater educational impact. It is also possible that VSTF texts can enhance students' homework reading assignments, without direct teacher supervision. All of these questions merit further empirical research.

The VSTF feature is now available as reading tool on electronic textbooks across every major subject area, from grades 3 through 12, by several US educational publishers, and on an expanding list of college electronic textbooks. In these high-volume, but less ephemeral educational publications, additional fabrication steps are needed to fully integrate the VSTF feature with other instructional and digital elements of the content. Dedicated software parsing engines and special fabrication tools have been developed to meet the special needs of this industry.

Also now available is a a user-directed, parse-on-the-fly capability, which provides access to parsing engines over the Internet, in real-time. Thin-client applications permit a reader to select several pages of text at a time, and receive back, within seconds a version of the text that has been transformed with the VSTF format. Derivatives of this parse-on-demand technology are now being adapted for special news and information web-based portals, and for internal use in large systems, such as legal and medical record systems, where digital text is used.

Future Research

Further research is also needed to better understand the mechanisms and optimal conditions by which the VSTF method can improve reading performance. How does the availability of VSTF text affect eye-movement, visual, lexical, syntactic, and discourse processes, and the dynamic collaborative interactions between these processes? At what age should VSTF optimally be introduced? How does VSTF interact with student improvements in fluency, syntactic awareness, and higher-level discourse processes? Does the improvement in reading comprehension endure without needing to continue using VSTF after the first year? Can further increases be seen if the method is used over more than one year? If the VSTF method enables students to read material that would have otherwise been too difficult for them without it, then, can this new, higher level of reading proficiency, in turn, enable them to attempt even more difficult material, for which the VST method would help them yet again?

The VSTF method is major departure from established text formatting conventions, including more recent methods based on digital typography [26] . It is the most fundamental change in written-language symbol arrays since medieval scribes inserted spaces between words more than 1000 years ago -- a change that also significantly improved text readability for many readers [27, 28]. Nevertheless, the present study demonstrates that the VSTF can be practically implemented, using computers that are becoming affordable for a large majority of the US population, and in actual classroom settings, not just the laboratory. Therefore, it can be expected that the VSTF method will be used across a wider range of texts than just textbooks in the classroom.

Moreover, reading remains a challenge for many adults, even for many who graduate from college. In a recent survey of adult literacy, the US Department of Education found that the proportion of *college graduates* who can read at a proficient level has dropped, in the past decade alone, by one-fourth, from 40 percent to 31 percent [29]. In addition, *more* people in the global economy are reading English as a *non*-native language than as a native language; even among those highly educated in English as a second language (e..g, Scandinavian physicians), reading comprehension diminishes by at least 25 percent compared to reading the same text in one's native language [30]. Because syntactic awareness is an important, independent predictor of sentence comprehension when reading in a non-native language, the visual-syntactic cues of VSTF can assist non-native readers in a task that directly augments sentence comprehension [31]. Finally, over half the sentences in the technical material used in today's information economy require complex syntactic structures, which challenge many readers, but which are refractory to editorial simplification [32, 33].

Therefore, the VSTF method may play a central role in improving reading performance across a wide range of ages and readers, as well as types of electronic text media. For these reasons, additional research on the basic neurophysiology of reading in VSTF will be of interest.

References

[1] Hauser, M.D., Chomsky, N., & Fitch, W.T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569–1580.

[2] Ferreira, F., & Anes, M. (1994). Why study spoken language? In (M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (ch. 2, pp. 34–36). San Diego, CA: Academic.

[3] Steinhauer, K. , Alter, K., & Friederici, A. D. (1999) Brain potentials indicate immediate use of prosodic cues in natural

speech processing. *Nature Neuroscience* vol. 2 (2): 191-196. February 1999.

[4] Cutler, A.D., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*, 141–201.

[5] Grodner, D. & Gibson, E. (2005). Consequences of the Serial Nature of Linguistic Input for Sentential Complexity. *Cognitive Science: A Multidisciplinary Journal.* Vol 29(2) Mar-Apr 2005, 261-290.

[6] Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. Journal of Experimental Psychology: Learning, Memory, and Cognition. Vol 16(3) May 1990, 417-429. Henderson, John M; Ferreira, Fernanda.

[7] Caplan D (2004). Functional Neuroimaging Studies of Written Sentence Comprehension. *Scientific Studies of Reading*, *8*(3), 225–240.

[8] Rayner, K.; Kambe, G.; Duffy, S. A. (2000) The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*. Vol 53A(4) Nov 2000, 1061-1080.

[9] Vidyasagar, T.R., & Pammer, K. (1999, April 26). Impaired visual search in dyslexia relates to the role of the magnocellular pathway in attention. *Neuroreport* 10(6), 1283-1287

[10] Horowitz, T.S., & Wolfe, J.M. (1998). Visual search has no memory. *Nature*, *394*, 505–507

[11] Rayner K, Chace KH, Slattery TJ, Ashby J (2006). Eye Movements as Reflections of Comprehension Processes in Reading. *Scientific Studies of Reading, 10*(3), 241–255

[12] Tabor, W & Tanenhaus, M. K.(1999) Dynamical models of sentence processing. *Cognitive Science: A Multidisciplinary Journal.* Vol 23(4) Sep-Oct 1999, 491-515.

[13] Bringuier V, Chavave F, Glaeser L, Fregnac Y, (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* 283: 695-699. 29 Jan 99.

[14] Engbert R, Nuthmann A, Richter EM, and Kliegl R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review,* 112 (4): 777–813

[15] Guri-Rozenblit, S (1988). Impact of diagrams on recalling sequential elements in expository texts. Reading Psychology. Vol 9(2) 1988, 121-139.

[16] Guri-Rozenblit, S (1989) Effects of a tree diagram on students' comprehension of main ideas in an expository text with multiple themes. *Reading Research Quarterly*. Vol 24(2) Spr 1989, 236-247.

[17] Moore, P.J. ; Chan, L. K; Au, W. K. (1993) High school students' use of diagrams during reading.. *Journal of Research in Reading*. Vol 16(1) Feb 1993, 57-71.

[18] Keller, T.A., Carpenter, P., Just, M.A. (2001). The neural bases of sentence comprehension: A fMRI examination of syntactic and lexical processing. *Cerebral Cortex*, *11*(3), 223–237.

[19] Hale, J. Uncertainty About the Rest of the Sentence. (2006) *Cognitive Science: A Multidisciplinary Journal*. Vol 30(4) Jul-Aug 2006, 643-672.

[20] McBride, R.(1976) Visual phrasing cues as an aid to comprehension of simple and complex sentences by learning disabled and normal pupils. *Journal of Research & Development in Education.* Vol 9(Mono) 1976, 109-110.

[21] Mokhtari, K. Thompson, H. B. (2006) How Problems of Reading Fluency and Comprehension Are Related to Difficulties in Syntactic Awareness Skills Among Fifth Graders. *Reading Research and Instruction.* Vol 46(1) Fal 2006, 73-94.

[22] Laberge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.

[23] Walker, S., Schloss, P., Fletcher, C.R., Vogel, C.A., & Walker, R.C. (2005, May/June). Visual-Syntactic Text Formatting: A new method to enhance online reading. *Reading Online, 8*(6). Available at: http://www.readingonline.org/articles/art_index.asp?HREF=r_walker/index.html
.

[24] Northwest Evaluation Association (2004). *Reliability and Validity Estimates: NWEA Achievement Level Tests and Measures of Academic Progress.* Northwest Evaluation Association. Lake Oswego, Oregon . Document available at: http://www.nwea.org/assessments/researchbased.asp

[25] Olejnik S & Algina J. (2000). Measures of Effect Size for Comparative Studies: Applications, Interpretations, and Limitations. *Contemporary Educational Psychology* 25**,** 241–286

[26] Knuth DE & Plass MF (1981). Breaking Paragraphs Into Lines. *Software — Practice and Experience* 11: 1119–1184 (1981). Reprinted as ch. 3 of *Digital Typography*, p. 67–155.

[27] Manguel, A. (1996). *A history of reading*. New York: Penguin.

[28] Saenger, P (1997) *Space Between Words: The Origins of Silent Reading.* Stanford University Press. Stanford, California.

[29] US Department of Education, National Center for Educational Statistics (2005). *The Health Literacy of America's Adults.* Available at: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006483

[30] Gulbrandsen, P, Schroeder, TV, Milerad, J, & Nylenna, M. (2002). Paper or screen, Mother Tongue or English: which is better? A randomised controlled trial. *JAMA*, 287(21), 2851-2853

[31] LeFrancois P and Armand F, (2003). The role of phonological and syntactic awareness in second-language reading: The

case of Spanish-speaking learners of French. *Reading and Writing: An Interdisciplinary Journal.* 16: 219–246.

[32] Hindle D.& Rooth M. (1993). Structural ambiguity and lexical relations. *Computational Linguistics. 19*, 103–120.

[33] Roberts, J.C., Fletcher, R. H., & Fletcher, S. W. (1994). Effects of peer review and editing on the readability of articles published in Annals of Internal Medicine *JAMA.* 1994;272:119-121

About the Authors

Randall C. Walker is an educator and clinician-scientist at Mayo Clinic College of Medicine. His undergraduate (Notre Dame) and medical (Mayo Medical School) training focused on linguistics, neurology, and internal medicine. His work includes the design and implementation of prospective, randomized, double-blind controlled trials of therapeutic interventions. Over the past 10 years, Randall has collaborated with his brother, Stan, to integrate current understandings in linguistic and visual processing of reading into a dynamic "cataphorical" model of spatial-syntactic and attentional processing. Randall developed the natural language parsing algorithms and Stan established the visual parameters, which were then combined and encoded by Phil Schloss into the visual-syntactic parsing engines used in this study.

Phil Schloss worked for 20 years as a senior software engineer and software system architect at IBM, where he patented 14 inventions, most entailing text-processing methods and Internet protocols, and published numerous technical disclosure bulletins and technical reports. He has also worked in development of natural language translation software and as the architect and lead developer of software development projects. In 1993, he received a special commendation from the governor of Minnesota, in recognition of his contributions to the economy and welfare of the state. He has worked on the visual-syntactic formatting method since 1996, including the engineering and design of computer-executed parsing engines and multidimensional text display methods. The research in the current article was performed using enhanced text that was the direct output of software parsing engines and display modules he developed.

Charles A. Vogel received his Ph.D. from the College of Education at Denver University; his dissertation research included classroom-based research with visual-syntactic formatting. He has over 25 years' classroom teaching experience in social sciences and language arts at the high school and middle school levels. For the past several years, he has been using visual-syntactic formatting in a wide range of classroom settings, involving several hundred high school students. His ongoing research with the visual-syntactic format includes a further analysis of its impact among non-native English language learners; the use of real-time parsing software by students as a tool to improve their own writing; and the use of visual-syntactic texts and software in preparing high school students for college-level reading.

Adam Gordon received his undergraduate degree from Yale University. After an extensive career in education and communication, followed by ventures in Internet-based technology development, he joined Walker Reading Technology in 2001, where he has focused on research management, and product and systems development. He works directly with Phil Schloss and Randall Walker in the ongoing development of VSTF-based parsing systems, inferface design, and network systems management.

Charles R. Fletcher received his doctorate from the University of Colorado at Boulder in cognitive science. His primary research efforts have been directed toward demonstrating experimentally the existence of three levels of representation in memory for discourse: a surface-level representation, a propositional textbase, and a situation model. This research into the online processing of discourse has focused on the role of memory and attention in understanding the causal structure of narratives; it involves both constructing computer models that simulate the flow of ideas through a reader's awareness during narrative comprehension and recall, and conducting experiments with human subjects to evaluate basic assumptions of those models. Recently, this research has included how syntactic and semantic factors interact to control a reader's attention.

Stan Walker is a vision scientist and clinical ophthalmologist involved in the education of medical students and graduate physicians, and in the practice of ophthalmology. As a practicing eye physician, he has evaluated and treated thousands of patients who have complained of visual reading problems, including eyestrain symptoms associated with reading from computer displays. A graduate of Notre Dame and Mayo Medical School, he prepared the visual rationale and the visual criteria that are used in the algorithmic generation of the visual-syntactic text format, and participated in the design of reading research reported in this article.