Visual-syntactic text formatting uses natural language processing techniques to parse sentences and present them in a way that highlights meaning. Specifically, VSTF breaks sentences up at salient clause and phrase boundaries, fits each row of text into one or two fixation eyespans, uses a cascading pattern to denote syntactic hierarchies, and creates visual clusters across multiple rows to help readers retain and integrate multi-phrase images. This paper reviews the theoretical basis of VSTF and summarizes research on its impact on reading comprehension, speed, retention, and proficiency, with a focus on its use by English language learners.

Introduction

The global volume of scientific output doubles approximately every 10 to 15 years and the vast majority of this international scholarship is published in English (see Crystal, 2003). At the same time, a high percentage of global commerce is also conducted in English, as are international debates and discussions in political and cultural realms. For all these reasons, the ability to read English is critical for participation in the knowledge economy and information society. Yet many countries, such as Japan, find the English reading ability of their citizenry lacking.

To help address this problem, second language researchers have long looked at the potential of digital tools to improve reading comprehension and proficiency (for an overview, see Chun, 2011). Forms of digital scaffolding previously explored for second language readers include first and second language vocabulary glosses (Yoshii, 2006);
visual, audio, and audio-visual supports (Chun & Plass, 2000); advance organizers (Len & Chun, 2007); highlighting of words (Hzang & Liou, 2007); text-to-speech (Roberts, Park, & Takahashi, 2010); and re-sequencing vocabulary introduction (Ryoo & Brown, 2008).

All of these approaches assume that text is presented in the same traditional block formatting in which it typically appears in books. This formatting is arbitrary, though, and is not necessarily the most advantageous for reading, especially among second language learners.

This paper discusses an alternate approach to digital scaffolding that involves a radical re-thinking of text formatting. We first present a brief history of text formatting and reading, then analyze what the challenges are to reading in traditional block formatting, next introduce a new approach called visual-syntactic text formatting (VSTF), and finally summarize research conducted on reading via VSTF.

**Brief history of formatting alphabetic texts**

The structure and formatting of texts has changed a number of times over the millennia. The Phoenician alphabet was written in 2000 BC without vowels, punctuation, lower case letters, or spaces between words (Penney, 2006), which, in contemporary English, would yield sentences such as this:

```
FRSCRNDVNYRSGRFTHRSGRHTFRTHNTHSCSCNTINENTNWNTIONCNCVEDNLBR
TYNDDCTDTTHPRDRSTTHNTHITLEMNRRTKTDQL
```

Vowels were added by the Greeks in about 1000 BC, and punctuation first appeared in Aristophanes’s plays about 200 BC (Fischer, 2001). It took nearly another millennium, until 700 AD, for lower case letters to emerge, which would render the above passage to appear like this:

```
Fourscoreandsevenyearsagoourfathersbroughtforth,uponthiscontinent,anewnation,
conceivedin Liberty, and dedicatedtothe propositionthatallmenarecreatedequal.
```

The above sentence is much easier to read, but still would be a challenge for anyone not already familiar with it.

A critical event in the development of literacy came with the insertion of spaces between words, which first emerged in Irish manuscripts in the seventh and eighth centuries and spread throughout Europe starting in about 900 AD, a change that would thus render the above sentence as follows:

```
Four score and seven years ago our fathers brought forth, upon this continent, a new
nation, conceived in Liberty, and dedicated to the proposition that all men are created
equal.
```

This relatively simple change had a profound effect on literacy practices. Prior to the use of spaces between words, almost all reading was done orally – either aloud, in groups, or individually, in a muffled voice – a necessary step to determine word boundaries and meanings. Word spacing facilitated the development of silent reading and thus, eventually, many of the practices we consider essential to modern literacy, such as the retrieval of reference information from a broad range of texts (Saenger, 2000).
Problems with block formatting

Sentence formatting has changed little since the tenth century, but the widespread shift of texts from paper to screen allows new possibilities previously impractical. Traditional block formatting of texts has two problems that can be addressed. First, due to limits of the human eye span, only nine to fifteen characters at average can be taken in at a time before moving to the next fixation when one reads conventional text (Demb, Boynton, & Heeger, 1997). Everything else on the page competes for visual attention as readers get lost in a sea of words (Walker, et al., 2007; see Figure 1). That is why even skilled readers have difficulty reading texts with wide blocks of texts and tiny margins, as opposed, for example, to texts that have been broken up into multiple columns.

In reading a block of text, when a person’s glance naturally shifts right to the next fixation, called a saccade, it sometimes requires regression to re-examine previously viewed words either due to interference with working memory (Garrod, 1992) or due to skipped words at the edge of a fixation (Rayner & Sereno, 1994). This may not only slow the speed of reading, but also impede reading comprehension, especially in the case of poor readers.

Skilled readers typically resolve this problem by grasping syntactic units in sentences, which helps them anticipate what should come next and how words, phrases, and clauses fit together in sentences to convey meaning. In oral speech, there are all sorts of clues to suggest sentence syntax, from pauses, to word stress, to changes in pitch, pacing, and flow. In written texts, however, most of these clues are missing, which can result in reader misunderstanding. Eye movement research suggests that skilled adult readers do not only activate a series of phonological segments, but also a prosodic structure during silent reading (Ashby, 2006).

All of these factors make reading much more difficult for second language learners, who, in some cases, lack explicit knowledge of the target language syntax, or, in other cases, may know the explicit rules but cannot rapidly and automatically use those rules to mentally process language. It is thus not surprising that knowledge of syntax has been
proven as one of the most salient factors in determining second language reading ability (see Shiotsu, 2010).

Reading comprehension is thus eased both through narrower columns and syntactically organized texts that assist readers in parsing syntactic meaning. This is seen both in computer programs, which are always written in a cascaded format that highlights syntactic structure and meaning, and in poetry, which often is.

**Visual-syntactic text formatting**

Over the last 15 years, a team of researchers has developed a new type of textual organization called visual-syntactic text formatting (VSTF). VSTF uses natural language processing techniques to automatically parse sentences and present them in a way that highlights meaning. Specifically, VSTF breaks sentences up at salient clause and phrase boundaries, fits each row of text into one or two fixation eye spans, uses a cascading pattern to denote syntactic hierarchies, and creates visual clusters across multiple rows that help readers retain and integrate multi-phrase images in their mind. VSTF also renders active verbs in colored font to further highlight meaning. The end result is a streamlined column of text that allows more efficient eye movement and syntactic processing. VSTF would thus render the above sentence from the Gettysburg address as seen in Figure 2 (with the highlighted words colorized instead of bolded):

![Figure 2. Visual-syntactic text formatting](image)

VSTF can be accessed by readers in three ways. First, a software product called ClipRead is available for free from Live Ink, the company that developed VSTF. With ClipRead open on a computer, students can copy and paste any digital text into the program for automatic and rapid conversion to VSTF, and then simply page through it. Secondly, instructors can prepare materials ahead of time for a class using the same ClipRead program. Third, a few publishers have begun to partner with Live Ink to have their digital textbooks pre-converted to VSTF, with readers of the digital texts having the option of shifting to reading the VSTF version simply by clicking on an icon, thus obviating the need for any selecting and copying of text.
Research on reading with VSTF

VSTF has been investigated in both laboratory and classroom settings, with junior high, high school, and college students, with very promising results (Vogel, 2002; Walker, Schloss, Fletcher, Vogel, & Walker, 2005; Walker & Vogel, 2005; Walker et al., 2007). Collectively, this body of research has addressed the following questions: (1) How does reading in VSTF as compared to normal block format affect text comprehension?; (2) How does reading in VSTF as compared to normal block format affect reading speed and efficiency?; (3) How does reading material in VSTF over the course of a year affect general reading proficiency (i.e., when measured by the ability to read material in normal block format)?; (4) How does reading in VSTF compared to block texts affect the user experience?; and (5) What are the particular effects of reading with VSTF on English learners? We will briefly summarize the methods, contexts, and results of key studies and interventions that relate to each of these five questions.

Reading comprehension

We have measured the impact on reading comprehension through several studies among college students, all of which have found a benefit for VSTF (see, e.g., Walker, Schloss, Fletcher, Vogel, & Walker, 2005). In an earlier study, 48 college students read six 500-word expository passages from a computer display: three in standard block format and three in VSTF. The order of passages and their format were randomized across participants. Immediately after reading the six passages, participants took a written test with 48 questions addressing the general nature of the texts, mid-level details, and specific details. A one-way ANOVA revealed that scores on the comprehension test were 40% higher for the passages read in VSTF as compared to those read in block format (p = .0024; see figure 3). Using students’ college admissions ACT test as a moderating variable, there was no significant impact of the ACT reading subsection scores on outcomes; in other words, both excellent readers and mediocre readers achieved the same benefit in comprehension from reading with VSTF.

![Figure 3. Comparison of comprehension test scores](image)
A more recent study confirmed these results, with an interesting added finding. In this study, we used a similar design to the above study, but in addition to VSTF and block format, we added a third condition, random truncation. In random truncation, lines were shortened to approximately the same length as that of VSTF, but they were not indented, and the line breaks were chosen randomly between words rather than by syntactic units. A total of 27 US college students participated in the study, and the reading passages consisted of science articles taken from an encyclopedia and of similar length. Comprehension questions included both text base questions, in which answers could be located within one or two adjacent sentences in the text, and situation model questions requiring integration of information from throughout the text.

Results indicated that there were no differences in comprehension between reading in block format and random truncation, but that the participants comprehended significantly better when reading VSTF than in either of these formats. In other words, narrower columns alone did not prove advantageous for comprehension, but were beneficial only when combined with cascaded syntactic formatting. On average, participants scored 20% higher on the comprehension tests in VSTF than in block format, with an effect size (Cohen’s d) of 0.25 (p<.01).

**Reading speed and efficiency**

In this recent study described above, the participants also wore eye-tracking equipment, which allowed us to measure the amount of total eye fixation time per word. Figure 4 provides a visual representation of a participant’s eye movement through two passages, one in block formatting and one in VSTF. As exemplified by this participant’s results, eye movement through the VSTF passages was much more streamlined and efficient.

![Figure 4. Eye movement in block text and VSTF](image)

Overall, participants spent an average of 780 milliseconds per word in reading the block text, but only 630 millisecond per word reading in VSTF, thus resulting in a 20% faster reading speed in VSTF (see Figure 5). This difference was not due to longer or shorter glances at each word, but rather due to more repetitive glances in the block text format as participants engaged in more frequent regressions. Participants carried out an average of
3.15 eye movements per word in the block text format, but only 2.25 eye movements per word in the latter format (see Figure 6).

![Figure 5. Average total fixation time (msec) per word](image1)

![Figure 6. Average regressive eye movements per word](image2)

Finally, the combination of comprehension data and reading speed data in the same study allows us to calculate the difference in reading efficiency (defined as amount of comprehension per unit of reading time) between the two main conditions. As the participants scored 25% higher in comprehension, and read at a 20% faster rate, we calculate the increase in reading efficiency due to use of VSTF at 50%.

**Reading retention**

Reading comprehension, speed, and efficiency are all of great value, but it is also important that students are able to retain information they read and thus learn academic content from reading. Retention has been investigated through two studies of high school students using VSTF to read content academic courses (Walker & Vogel, 2005; Walker et al., 2007). Both showed that students retained information better in VSTF.
For example, in 2003–2004, a study was carried out in six Grade 10 World History Classes, with three teachers each teaching two of the classes. First, the 84 students in the classes were randomly assigned to one of the six classes. Then, the three teachers were randomly assigned to teach one experimental course and one control course. A pretest showed that the students in the experimental and control groups were equivalent in academic ability, and the two groups also had similar proportions of gender and native language subgroups.

The history textbook used in the course was prepared for digital display and converted to VSTF. In each of the six classes, students read the textbook material on computer in twice a week 25-minute sessions at a computer lab. The experimental group read the material in VSTF, whereas the control group read the material in block format. Both VSTF and block text electronic textbook platforms permitted font enlargement, choice of dark or light background colors, and a 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.

Other than the textbook reading condition, all other instruction for the two groups was equivalent. Students had the same assignments and took the same unit exams (ten) and final exam (one) throughout the academic year. All of the exams were generated automatically by test-generator software and a question bank that the publisher provided with the textbook. The unit exam covered material taught over the previous few weeks whereas the final exam covered material taught over the academic year.

Exam scores of students in the experimental (VSTF) and control (block format) groups were analyzed with a test of repeated measures multivariate analysis of variance, with pretest reading comprehension as an independent variable. Results showed that students in the VSTF classes scored significantly higher on both the unit exams and the final exam.

Figure 7 shows the differences between the VSTF group and the block format group on the ten unit exams. Students reading the textbook with VSTF scored higher on each of the ten tests. In addition, there appears to have been a training effect, as the average gap between the scores of the VSTF group and the block format group was much higher in three of the last four weeks than in had been in any of the first six weeks. Overall, the effect size of the difference in exams in the second half of the year (.55) was larger than the effect size of the differences in the first half (.375).
Students in the VSTF group also scored significantly and substantially higher than those in the block format group on the final exam. The mean final exam score was 81 in the VSTF group and 64 in the block format group, representing a substantial difference. Figure 8 graphically shows the range of final exam scores of students in both the VSTF group and the control group. The graph shows that VSTF was advantageous for outcomes of both the better performing students and the worse performing students, but the overall benefit was greater for the latter as seen in the narrower bell curve of the VSTF group. Specifically, the highest performing students in the VSTF group scored approximately 9 points higher than their counterparts in the control group, whereas the lower performing students in the VSTF group performed more than 20 points higher than their counterparts. This difference can likely be attributed to a ceiling affect, as the highest achieving students in the VSTF group reached the maximum score of 100. Overall, as seen in the figure, the lowest performing students in the VSTF group almost reached the mean score on the final exam of the students in the control group.

![Figure 8. 10th grade final exam (mean +/-2 SD)](image)

**Reading proficiency**

The benefits of VSTF for reading comprehension and retention are not surprising given what we know about how the eyes and mind process information. The thornier question is the effect of VSTF use on long-term reading proficiency. Simply put, do students who read regularly with VSTF transfer skills learned and thus become more proficient readers of material in traditional block format? Or do they fail to do so, and thus suffer in reading proficiency because they have become dependent on the scaffolding provided by VSTF and less able to read well without it?

This question has been examined through multiple studies of students in middle and high schools (Vogel, 2002; Walker, Schloss, Fletcher, Vogel, & Walker, 2005; Walker & Vogel, 2005; Walker et al., 2007). In all studies, students who read textbook material in VSTF for 50 minutes per week over the course of a year were compared to students who read the same textbook material in traditional block format. However, two different study designs were used. In two high school studies, 9th and 10th grade students in both the VSTF and block format groups read social studies textbook material digitally, exactly as described in the study of retention above (for details, see Walker et al., 2007).
In contrast, in the middle school study, a comparison was made between reading VSTF on computer and reading the same material in block format in ordinary printed textbooks. In this study, 6th, 7th, and 8th grade students in the VSTF groups read social studies or language arts material on laptops in their classrooms for a total of about 50 minutes per week. Block format students were selected from elsewhere in the district based on matched case control: for each student in the VSTF group, a counterpart student was selected who had comparable attributes in terms of grade, baseline reading scores, English language learner status, and gender. The block format students read the same social studies and language arts material, but out of their ordinary printed textbooks. As the block format students were spread out in classrooms across the school district, there was no attempt to control the amount of time they spent reading the textbook material.

A total of 528 students participated in the studies, 384 in the middle school grades (6, 7, and 8) and 144 in the high school grades (9 and 10). In all grade levels, reading proficiency was measured by scores on the reading section of the Measure of Academic Progress Test by the Northwest Educational Association (NWEA), a standardized test given in much of the United States. On this test, all reading comprehension test passages are formatted in conventional block text, and participants in the study took the tests in the ordinary manner as provided by the test publisher. Changes in reading proficiency were indicated by the relationship between the NWEA pretest given at the beginning of the school year in fall and the NWEA posttest given at the end of the school year in spring. Pretest-posttest comparisons were performed with an analysis of covariance, using pretest reading scores as the independent variable and posttest reading scores as the dependent variable.

In all five grade levels from 6th to 10th, students in the VSTF groups made significantly greater improvements in their reading proficiency over the course of the year than did their counterparts in the block format condition (see Figure 9). Students in the block text formats typically made small gains throughout the year, advancing less than one grade level when compared to national averages. In contrast, 6th, 7th, and 8th grade students reading in VSTF made approximately two years’ worth of gains in a single year. Gains for 9th and 10th grade students in the VSTF groups were also impressive; however, it is not possible to quantify them in terms of years of growth since the NWEA reading test is not given or normed past the 10th grade.

The positive impact on reading proficiency through use of VSTF may appear counterintuitive as the testing measures used involved texts written in regular block format that the participants had never seen. However, there are theoretical bases for understanding how this transference from reading in VSTF to better reading in block text format might occur. Knowledge of syntax is shown to be an important precursor to reading ability, as is knowledge of English prosody (Korlat, Greenberg, & Kreiner, 2002). Other interventions, such as reading aloud to children, that help learners understand the prosody of structure of texts have similarly been shown to improve reading ability (Dowhower, 1991). The positive impact of VSTF use on general reading proficiency may be due to this training effect, as readers gain a better understanding of how written passages are structured and are thus better able to subconsciously parse sentences they read. Additionally, there may also be a confidence-building effect; as students realize they can successfully read and understand material, they may approach new texts, in any format, with less fear or disinterest.
Surveys have been carried out of both high school and college students as to their preference to reading in VSTF or ordinary block formatting on the screen. All surveys show strong preference for reading with VSTF. For example, in a survey of college students after one hour of use, 60% indicated they found it easier to read in VSTF compared to only 20% who found reading easier in block text format (Walker & Vogel, 2005). Some 41% of college students who read block text on the screen for an hour reported eye strain, compared to 11% of those who read VSTF on the screen for the same amount of time, representing a 73% reduction in reported eyestrain (Walker et al., 2005).

Among high school students who used VSTF over the course of an entire school year, the preference for the cascaded format was even greater. Some 85% of students indicated that they preferred VSTF whereas only 9.5% said they preferred to read digital material in block text format. High school students in the studies described previously were also given the choice of abandoning digital reading during the year and reverting to their textbooks. No students in the VSTF groups chose to do so.

A study that we are currently carrying out in 50 4th and 6th grade classrooms in California, many with large numbers of English language learners, has provided further data on learners’ experience with VSTF. All these classrooms are part of one-to-one laptop programs, where students have access to an individual laptop computer throughout the school day. At the time of this writing, we have completed approximately 30 observations in classrooms using VSTF, and the students and teachers we have interacted with have been almost universal in their enthusiasm in reading with VSTF. One student told us, “It is easier, faster and more fun to read with it because words are easier to see and the sentences are broken up.” Another student said, “It is not only easy and fast, but also has more information.” Although she did not elaborate on what additional information VSTF carries to readers, it seems that she could distinguish more details and make more sense of what she read due to the way VSTF texts are parsed. Teachers were especially gratified to see how use of VSTF texts boosted slow readers’ confidence in reading. Teachers have also told us...
that their students began reading in phrases when using VSTF, rather than just sounding out each word as they typically did when reading block texts. Some of the teachers who have students with special needs, such as autism or vision problems, suggested that this particular format has worked well for such students, saying a laptop with VSTF texts is like an “added bonus.” Here is a dramatic example of this: According to a teacher, one student, who has an eye focusing and tracking problem, one day went home and said, “Mom, I can read like a normal person!”

**Effects with English language learners**

All of the above has referred to the general effects of reading in VSTF. What then do we know about the particular effects of VSTF reading among English language learners? In a year-long study with 9th grade students, greater gains were seen with lower aptitude students than among those who were already good readers (Walker et al., 2005). They found that many of these students were English language learners who were able to achieve the same reading proficiency level at the end of the year as the control group of native English speaking readers. Walker and Vogel (2005) reported that the use of VSTF had a significant positive effect on English language learners’ reading development across all secondary grades, though, compared to their native peers, it took more sessions until their quiz scores got an increase over control students. By the end of the year, in most grades, English Language learners in the VSTF cohort had closed one-half to nearly the full gap between themselves and the native English students in the control groups.

How does this particular effect relate to what we know about second language reading in English? Among native English speakers, the size of one’s vocabulary is viewed as more important than syntax in affecting reading ability, because vocabulary knowledge in one’s native language varies greatly while knowledge of syntax is relatively invariant (Shiotsu, 2010). However, among English language learners, knowledge of syntax has been found to have at least as great an effect, and in many cases, a greater effect on reading ability than does vocabulary size (see Shiotsu, 2010). Functional brain image studies also show that the patterns of brain activity when one reads a foreign language are fundamentally different from those when reading one’s native language (Dehaene et al., 1997; Kim, Relkin, Lee, & Hirsch, 1997). Similarly, a behavioral study demonstrated that the more dissimilar the syntactic rules of a target language are from one’s native language, the more difficult reading in the target language becomes (Frenck-Mestre & Pynte, 1997). All this helps explain why English learners in particular benefit from reading syntactically-parsed texts and why they attain a training effect from doing so.

This has been witnessed in schools and colleges. For example, Fisher and Newbury (2009) report that English learners at their community college find use of VSTF especially helpful, as it simplifies the reading process and makes complex passages easier to understand. Recently, a high school in Colorado has implemented a program for standardized test preparation for low-performing students, two-thirds of whom are English language learners. To prepare for the Colorado student assessment program (CSAP), the students participate in 20-minute sessions on a daily basis for four weeks, reading sample test passages in VSTF. All of the students, who begin the preparation program at an unsatisfactory or partially proficient level, are able to improve their reading proficiency (Vogel, 2011). Most recently, 81% of the participants met the Colorado state standard for acceptable growth in one year after a four weeks of 20-minute per day sessions, and 62% met the school goal of reading
at a proficient level by the end of the school year. Referring to the name of the software company that developed VSTF and has become synonymous with it, Vogel explained why he thinks it has been so successful in improving the reading proficiency of English learners and other struggling readers:

With reading, students in the first three grades develop word sense. Where we lose kids after the third grade is with sentence and paragraph sense. Comprehension is in the syntax. Kids get lost on the “garden path” because they can’t organize phrases and clauses into meaningful ideas. They re-read and waste time, becoming more frustrated. After a while, they forage instead of read. We have responded to these problems by abridging texts, in effect dumbing down the content. The best students breeze through these textbooks. It is an easy task, an easy “A.” The problem arises, when those brains, conditioned to read simple syntax, are challenged. I see it every year....

One question always arises, “Well, kids can’t read from Live Ink all the time! What about that?” I would suggest that there is a transfer effect to practicing Live Ink reading. The CSAP tests that those students passed to proficiency were not in Live Ink. The ACT test, all the assessments I have tracked students by were not in the format, and there have been good solid gains in comprehension with these students. Anecdotally, students will tell me that they can “pull out the phrases” and “see the verb” easier with other reading after working with the format.

Conclusion

There have many efforts to develop more effective techniques, approaches, and interventions to aid second language reading, including better uses of technology. Nevertheless, at least in some countries, reading skills in English as a foreign language seem to have stagnated. To address this situation, it is probably helpful to look at reading and reading education in fresh ways. We have long taken reading in block format for granted, yet its arbitrary structure does not convey as much useful information for comprehension as spoken language does. In contrast, considerable evidence suggests that VSTF provides valuable support, not only for strengthening comprehension of the text at hand, but also for improving general reading proficiency. More expansive research on this promising approach is called for.

References


Roberts, K., Park, H. J., & Takahashi, K. (2010). Effectiveness of text-to-speech software features for improving vocabulary, reading comprehension, and reading rate of culturally diverse high school students in Hawai‘i. In D. Gibson & B. Dodge (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010* (pp. 1606–1611).


Author biodata

Mark Warschauer is Professor of Education and Informatics at the University of California, Irvine, and editor of Language Learning & Technology journal. His most recent book is Learning in the Cloud: How (and Why) to Transform Schools with Digital Media.

Youngmin Park is a PhD student in Education at the University of California, Irvine, specializing in Language, Literacy, and Technology. Previously a high school teacher and teacher trainer in Korea, her research focuses on the use of digital media in the teaching and learning of reading.

Randall Walker, MD is an assistant professor of medicine at Mayo Clinic. His research applies natural language processing to improve text readability for persons with low-vision and reading disabilities, and has received awards for innovation and distinguished achievement from the US Department of Education and the Association of Educational Publishers.