

The effects of syntactically parsed text formats on intensive reading in EFL

John C. Herbert

National Institute of Technology, Akashi College, Japan
herbert@akashi.ac.jp

Separating text into meaningful language chunks, as with visual-syntactic text formatting, helps readers to process text more easily and language learners to recognize grammar and syntax patterns more quickly. Evidence of this exists in studies on native and non-native English speakers. However, recent studies question the roll of VSTF in certain EFL contexts. The report herein shows eye tracking data and numerical results from EFL students engaged in intensive reading activities, which utilized parsed texts or conventional block texts in controlled experiments. This report helps readers to decipher whether or not derivatives of VSTF are appropriate for their specific contexts.

Keywords: VSTF, Parsed Text, Scanning, Skimming, Intensive Reading

Introduction

Intensive reading instruction in EFL, particularly in Japanese high schools and colleges, is often delivered to mixed level classes. This creates a dilemma in which the instructor must find a way to capture every student's eyes literally on the same page regardless of each student's reading comprehension level. If the reading is too easy, the higher level students could lose interest quickly. Likewise, if the reading is too hard, the lower level students might just give up on reading it. Therefore, the type of reading task and the format in which the text is delivered become crucial for achieving an educational experience which is more balanced for all of the learners.

Regarding the task, higher level reading ability does not necessarily equate to a better performance on activities such as timed **237**

Regular Paper

scanning and skimming. At the same time, lower level reading ability should not imply that one cannot be taught to scan and skim challenging texts. Instead, through timed scanning and skimming, the possibility that a lower level reader might happen to come across key information that a higher level reader might have skipped over in the same amount of time increases. Therefore, scanning and skimming integrated with other group reading and discussion activities have played an important role in the author's mixed level EFL classrooms. (See Herbert, Herbert, & Bergman, 2003.)

Furthermore, there is an array of ideas out there concerning the manipulation of text formats as a means to make a reading more accessible for lower level readers. For example, children's books utilize larger fonts and considerable white space. There are fonts that target readers with dyslexia. In business English, bullet points, lists, and charts make it easier to grasp key concepts. All of these ideas could be incorporated into EFL textbooks by a publisher or in online or electronically generated readings manipulated by teachers where allowed. And, then there is the e-book. Too small? Zoom in. Don't know the word? Just tap on it. How much easier will this be in ten years from now? Hopefully, it will be much easier. There is still at least one more big change e-books have not embraced. That is, if a publisher were to create a means for the reader to simply push a button and convert the text format from block to syntactically parsed, as in the manner described hereafter, the reader could potentially have a less challenging reading experience.

Background

At the JALTCALL 2011 Conference in Kurume, Dr. Mark Warschauer's plenary address stimulated conference participants' interests in visual-syntactic text formatting (VSTF), which is a form of text manipulation involving syntactically parsed chunks. Dr. Warschauer portrayed VSTF as the next logical step in the evolution of written text on the basis of the format being much easier to read because of the way it requires the visual parsing of text into meaningful language chunks which cascade down a page with verbs indicated by color changes (Warschauer, 2011, June). (See Figure 1.) The reformatting of conventional text into VSTF has proven to be beneficial for native and non-native English readers across the United States with respect to reading efficiency and comprehension in a number of studies and surveys (Walker, Schloss, Fletcher, Vogel, & Walker, 2005; Walker et al., 2007; Warschauer, Park, & Walker, 2011). Dr. Warschauer's talk and the aforementioned literature have provided a catalyst for the study herein. Henceforth, this report examines whether or not derivatives of VSTF are appropriate for EFL contexts with a particular focus on intensive reading.

Previous research in support of VSTF

When VSTF was introduced at the National Educational Computing Conference 2005 in Philadelphia, five unique features that define the format were explained as follows:

1. Lines break at phrase and clause boundaries
2. Shorter rows of text fit in one or two fixation spans
3. Cascading depicts syntactic hierarchies
4. Row-clusters remain vivid in "mind's eye"
5. Indentations guide the eyes from row to row (Walker & Vogel, 2005, p. 4).

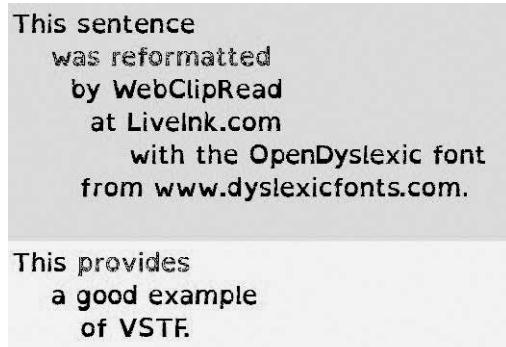


Figure 1. Sample text presented in a visual-syntactic text format
(Walker Reading Technologies, Inc., 2014).

Here one fixation span, or eye span, could be defined in accordance with Demb, Boynton, and Heeger (1997, as cited in Warschauer, Park, & Walker, 2011, p. 257) to be nine to fifteen characters in length. Therefore, any line with 30 letters and spaces could be considered undesirable for this format.

Regarding a study involving 48 college students who read three computer-displayed, 500-word texts in a block format and three more in a VSTF format matching that described above, comprehension of the VSTF texts was 40% better than that of the block texts. However, VSTF was read 10% more slowly. The researchers suggested that the extended time on task could have been attributed to the fact that their participants reported less eye strain from looking at the computer display when reading the VSTF texts. They also claimed that, even with the slower reading rate, "reading efficiency (comprehension score divided by passage reading time) was 25% higher with the visual-syntactic format" (Walker, Schloss, Fletcher, Vogel, & Walker, 2005, p.8).

Later, Walker et al. (2007) focused on the reading comprehension scores of 10th and 11th grade high school students in the US reading electronic textbooks twice a week for one academic year during their history classes. The control group read their history books in a block text format while the experimental group read theirs in VSTF. The gain scores in reading comprehension as measured by two versions of a standardize test used as pre- and post-tests were significantly greater for the experimental group than the control group. Both groups consisted of native English speakers and ESL students; however, the researchers admitted that most of the ESL students had been in the US for a fair amount of time, came from Spanish speaking backgrounds, and had good commands of English (p. 7).

At some point, a sixth feature was added to the aforementioned criteria for defining VSTF. That is, in more recent reports on VSTF related research and as part of the VSTF conversion software, WebClipRead by Walker Reading Technologies, Inc., colored fonts for active verbs were introduced "to further highlight meaning" (Warschauer, Park, & Walker, 2011, p. 258). In addition, WebClipRead now includes an optional OpenDyslexic font for making the format even more accessible for readers with dyslexia.

However, it is important to note the following points about the WebClipRead software, before putting it to use. According to Walker, Schloss, Fletcher, Vogel, and Walker (2005, 239

p.4). "For professionally composed, expository English writing (e.g., technical and business communication), this software is consistently able to differentiate principal phrases and specify the hierarchical relations between phrases for most sentences. When writing narrative, personal, or more stylized, phrase groups can still be identified, but clause structure extraction is more difficult." In other words, optimal parsing is not always achieved for each writing style with WebClipRead. Therefore, researchers, teachers, and readers must bear this in mind when applying VSTF to text via WebClipRead. It is for this reason that the derivative of VSTF employed herein has not been generated from WebClipRead, but manually by the researcher instead. As discovered in the early stages of this project, even adapted expository prose for EFL learners is likely not to fit into the "professionally composed," "technical and business communication" style desired for using WebClipRead. Therefore, when formatting a text with VSTF, EFL teachers must make sound pedagogical decisions on their own about what constitutes an appropriate boundary for a phrase or language chunk based on the target audience's language knowledge and learning needs.

Previous research less supportive of VSTF

A review of related literature, outside of those above which advocate the use of VSTF, led to two research reports, each worthy of its own set of disclaimers, but both with claims that VSTF might not be as uniformly helpful as originally thought for Asian learners of English in the context of their reading research projects. Researchers working with 30 mostly Chinese and other East Asian students at the Massachusetts Institute of Technology compared reading efficiencies observed when using a standard block format, VSTF, and a format they referred to as the Jenga format. The Jenga format simply involved the adding of white space between each complete sentence of a paragraph, and it yielded significantly higher gains in "comprehension per unit time" over VSTF but not over the block format. In addition, the researchers found mixed reactions from participant interviews regarding VSTF. Some preferred the syntactic parsing and colorization of verbs while others complained about the number of line breaks per sentence. Furthermore, the same participants had varied performance results which attested to VSTF actually slowing down the reader with no significant effect on comprehension scores (Yu & Miller, 2010). Later, a researcher at Tokyo Metropolitan University applied VSTF to a software application referred to as Multimedia Player Mint, which was designed to help students read faster. He concluded that VSTF did enhance the English reading comprehension and reading efficiency of 17 Japanese readers, but it could not increase the participants reading speed itself to any significant degree in the context of that application (Kanda, 2012).

The Mint application reveals text to a reader one chunk at a time, adding to an incomplete paragraph piece by piece until the paragraph is full. Each newly added chunk is presented in specific breath groups which are underlined to capture the reader's eyes and encourage faster, forward reading. However, whereas the Mint application appears to work better when meaningful language chunks are defined as subjects, predicates, or complete clauses, that constitute a reason to pause, VSTF is noted for shorter phrases which can fit into fewer eye spans per line. Therefore, the researcher attributes the failure of VSTF to improve reading speed significantly in the Mint application to the increased frequency of line breaks which interrupt the otherwise smoother subvocalization of text (Kanda, 2012).

Rationale

In the empirical study described herein, the participants were required to complete 20 intensive reading assignments as part of a control group or an experimental group over a three month period before taking a post-test in front of an eye-tracking device. The control group was trained to complete the post-test by carrying out 20 assignments in a block text format that were similar in readability and instruction to that of the post-test. During the same period, the experimental group had the same amount of training for the post-test with identical content and instruction, but the format of the content had been changed to a derivative of VSTF. Despite the slowed reading rates reported above, the intensive nature of the timed reading activities in this study accompanied by the parsed text format were expected to result in more productive scanning and skimming, more accurate topic guessing upon scanning and skimming, and faster reading with better reading comprehension through the prolonged exposure to the treatment. The goal of the intensive reading activities themselves was to improve these skills, and including parsed text in the mix of activities was expected to enhance these improvements.

In order to determine actual gains, numerical results are examined on the number of scanning words found under a time constraint; the accuracy of guessing the gist of a reading just by scanning, skimming, or reading; reading speed; and, comprehension as effected by both conventional block text and parsed text. Furthermore, this report includes heat maps displaying where the research participants' eye gaze plots indicate common fixation points while the participants scanned, skimmed, and read each of two final texts – one block and one parsed. These heat maps are important because they provide visual representations of how the readers' eyes collected information with very different eye movement patterns between each text format. Where the parsed text is used, the eye movement patterns reveal fixation points centralized within each language chunk while block text lends itself to more word-by-word fixation points. Advocates of VSTF believe this unique difference encourages readers to process language chunk-by-chunk rather than one word at a time; and, in turn, this leads to better reading efficiency (Warschauer, 2011, June).

Research questions

This study focused on the following research questions: In an intensive EFL reading context, does a syntactically parsed text format affect the . . .

1. Number of words a reader can find during a timed scanning drill?
2. Extent of one's ability to capture the gist of a reading during scanning, skimming, and timed reading?
3. Speed of reading upon completion of scanning and skimming drills of the same content?
4. Extent of reading comprehension?
5. Eye movements of readers during scanning, skimming, and timed reading?

Creating the parsed format

First and foremost, VSTF is a creation of Walker Reading Technologies, Inc. and the research teams behind LiveInk.com. Therefore, as that creation is produced by online software which is neither owned nor used by the author, there is no liberty to say that the parsed format **241**

utilized in this project is, in fact, VSTF. Instead, by applying one's own interpretation of how texts should be parsed, anyone can create a derivative of VSTF manually on their own. The derivative of VSTF created for this research project differed from that produced by WebClipRead in that the verbs were not colorized, and the determination of where to cut each chunk was decided by the researcher. The criteria used for determining how to divide up the language chunks included the following considerations:

1. Number of eye spans required for perusing one line of text
2. Natural pauses
3. Syntax familiarity

Following the aforementioned model set by Walker and Vogel (2005), one or two eye spans, or a total of up to 30 characters per line, was regarded as an optimal amount of text per line to minimize distractions from other surrounding text. Next, in an effort to minimize the interruptions in subvocalization due to frequent line breaks (Kanda, 2012), back indentation was used to indicate natural pauses, or new breath groups. In addition, one blank line was inserted between each sentence. But, there were no paragraph markers. Finally, with respect to syntax familiarity, teacher intuition was used to cater to the participants' presumed syntax knowledge. This was done either by adding or removing a line as needed without changing the content of the sentence itself.

Materials

The 20 readings which made up a prerequisite for participating in this project were selected from McGraw-Hill's SRA Reading Lab 1b. Ten of these readings were uploaded to a Moodle 1.9 LMS with two versions each, block and parsed, for the respective groups to complete as homework. The other ten readings were printed as classroom materials in the same two formats.

The reading materials used for the post-test were two 185 word expository texts about birds also from McGraw-Hill's SRA Reading Lab 1b, which had been slightly altered with permission, both in content and format. The minor content changes ensured the equal length and first grade readability level of each text.

Research participants

In order to recruit a large number of research participants, the 20 aforementioned reading activities were integrated into the curriculum of six third- and fourth-year English conversation classes at a five-year technical college in Japan. As such, the students ranged from age 17 to 21. One set of three classes had 118 regularly attending students, most of whom completed all of the reading assignments in a conventional block text format. Students pulled from these classes to take the post-test were recruited by monetary incentive for the project control group and referred to as the block text group, or BT group because their training in intensive reading focused on block formats. At the same time, the other three classes consisted of 113 regularly attending students, most of whom completed all of the reading assignments in a parsed text format. Therefore, students pulled from these classes were recruited, also by monetary incentive, for the experimental group and were referred

to as the parsed text group, or PT group because their training in intensive reading focused on parsed text. More specifically, their training focused on a derivative of VSTF.

All 231 of the English conversation students in regular class attendance responded to an informal survey about the parsed text format. And, the students, who completed all 20 assignments properly, were invited to participate in the study, which focused on two final readings, one block and one parsed. In total, 60 students were given monetary incentives in the form of ¥1,000 book coupons (or "toshō cards") to complete the final activities in front of an eye-tracking device and to disclose their collected data for research and publication purposes. There was no connection between the students' grades and these last two reading activities.

The first 58 participants stepped forward voluntarily. There were 28 who joined the BT group and 30 for the PT group. At least 100 other students qualified to participate by completing the twenty assignments, but they chose not to do so because of schedule conflicts, a lack of interest, or both. In order to balance the number of participants, two more BT group members were given personal invitations to participate in the experiment. Then, the 60 research participants were equally divided into two groups of 30 with 15 BT group members and 15 PT group members in each group; hereafter referred to as Group One and Group Two.

Procedures

In preparation for this research, the author carefully edited and parsed 185 words of *Dancers in the wild* (Vendetti, 2005) and *Who uses tools?* (Churchill, 2005) using the criterion discussed above without the aid of WebClipRead. Then, two tests were created with the adapted materials and entered into the EyeMetrix Design software program for eye-tracking. One test displayed the block text of *Dancers in the wild* and the parsed text of *Who uses tools?* for Group One. Group Two, on the other hand, was shown the block text of *Who uses tools?* and the parsed text of *Dancers in the wild*. Each reading was displayed in its entirety within a single view of a 21.5 inch monitor screen. The eye tracking equipment itself, which can be seen in the middle of Figure 2 below, was a non-obtrusive desk-mounted device that could be adjusted to accommodate the participant's physical position. Once the device was calibrated to track the participant's eye movements, that participant was asked to hold that position in a sturdy straight-backed chair during each individual task, but given freedom of movement between tasks without having to re-calibrate for each new task. This was very convenient for research purposes because there were no head supports or head-mounted cameras that might have otherwise obstructed the researcher's attempt at creating an ordinary reading experience.

The instructional procedures followed in this experiment mirrored the 20 reading activities the participants had completed for their classes in every way, except that the steps were performed in front of eye-tracking equipment in the researcher's office, both text formats were employed, and the participants did not discuss their topic guesses with classmates. The researcher followed these steps:

1. Demonstrate how to use the eye-tracking equipment with self-limiting head movements as allowed by the Mirametrix S2 Eye Tracker for up to $25 \times 11 \times 30\text{cm}$ (width \times height \times depth) (Mirametrix, 2014)
2. Discuss ten key vocabulary from each reading and confirm the participants' understanding of every word

3. Calibrate the eye-tracker with the participant's eye movements
4. Have the participant repeat the keywords out loud
5. Say each of the ten keywords out loud in 4 second intervals and in the order that they appear in the time-controlled text display, while the participant tries to left-click on each spoken word during eye tracking
6. Ask the participant what they think the reading is about
7. Display the text for 15 seconds for the participant to skim, not read, while tracking their eye movements
8. Ask the participant what they think the reading is about again
9. Time and track the participants reading of every word from start to finish
10. Ask the participant what they think the reading is about one last time
11. Have the students answer five comprehension questions
12. Repeat steps 4–11 with a second reading that is formatted differently.

Numerical results

Numerical results for the number of keywords found, the accuracy of topic guesses, reading speed, and comprehension scores were analyzed via t-tests for 49 research participants using SPSS statistical analysis software. The other 11 participants, who were excluded from the statistical analysis, experienced distractions such as computer malfunctions and loud noise from the school halls. As compensation was fixed for one hour of their time and that



hour had disappeared, their data was thrown out rather than replaced. This left 24 PT participants and 25 BT participants in the data pool. In each statistical analysis, the software's default probability value for keeping a null hypothesis, $\alpha = 0.05$, was used to calculate the critical value of t for a two-tailed t-test at $t = 2.0106$. For the true t -values calculated from post-test data in each of the four categories identified above, an absolute value of $t(48) > 2.0106$ and a probability of $p < 0.05$ were sought for throwing out each null hypothesis. However, the only category in which this degree of statistical significance was achieved was in the number of keywords found during the scanning task when the text was parsed.

Scanning results

Regarding the scanning activity, many of the research participants commented that the parsed text made it much easier for them to find the words that they were scanning the text to find. On average, the 49 participants found 8.53 out of ten words in each of the two parsed texts, while they only found 7.57 out of ten words within the block texts. (See Table 1.) As stated above, this discovery was the only one in the study which yielded statistically significant results based on a two-tailed t-test, where $t(48) = -2.6172$, $p = 0.0118$. As such, the researcher could claim with nearly 99% confidence that students, who have received training in scanning drills, should be able to scan for words more easily if the text were parsed. However, despite these encouraging results, a closer look revealed a bit of irony. That is, while 37 of the 49 participants found the same number of words or more when the text was parsed than when it was not, most of the participants who did not find the same number or more were the ones who had been in the PT group. Specifically, seven participants from the PT group (30% of all the PT group's research participants) found more words in the block text than the parsed text. At the same time, only five participants from the BT group (just 20% of all the BT group's research participants) found more words in the block text than the parsed text.

Table 1: Statistical analysis on scanning results (number of words found)

$\alpha = 0.05$	Critical value: $t = 2.0106$ (two-tailed)	
	Block	Parsed
Average =	7.57	8.53
Std. Dev. =	2.14	1.50
Higher score distribution	12 (7) > =10 (4)=	< 27 (13)
All (PT only)		
	$t(48) = -2.6172$, $p = 0.0118$	
	Interpretation: Throw out the null hypothesis	

Topic guessing results

Another noteworthy result about introducing parsed text to the scanning exercise was that it led to more accurate topic guesses just by quick glances over the reading content. Normally, one might not imagine using a scanning activity as a comprehension measure. The purpose of scanning is not to find the same general information that skimming helps one find, but to find specific details. However, throughout this project, the keywords that 245

were scanned for and any information that may have stuck with the participant after catching his or her eyes during the scanning task were used by participants to guess the gist of the reading. Only seven out of 49 participants could guess the main topic of the reading by scanning the block text, but 13 out of 49 guessed the topic correctly after scanning the parsed text. (See Table 2.) After skimming and timed reading, the number of additional correct topic guesses went up steadily; but, by the end of the activity, there were still 14 of 49 participants who could not understand the block text reading, while only eight participants could not understand the point of the parsed text. Of the twelve participants who found it easier to guess the topic of the block text reading, four were from the PT group. Based on the two-tailed t-test for topic guessing, where $t(48) = 1.9415$, $p = 0.0581$, the numbers fall just short of the necessary t-value and probability for claiming a statistical significance. (See Table 3.) Therefore, though noteworthy, the researcher must concede that these results are not widely generalizable.

Table 2: Number of correct topic guesses after each task

	Block	Parsed
After Scanning (3pt)	= 7/49	13/49
After Skimming (2pt)	= 21/49 (+14)	26/49 (+13)
After Reading (1pt)	= 35/49 (+14)	41/49 (+15)
Incorrect (0pt)	= 14/49	8/49

Table 3: Statistical analysis on topic guesses

$\alpha = 0.05$	Critical value: $t = 2.0106$ (two-tailed)	
Block		Parsed
Average =	1.29	1.63
Std. Dev. =	1.04	1.05
Higher score distribution	12 (4) >	=13 (8) = < 24 (12)
All (PT only)		
$t(48) = 1.9415$, $p = 0.0581$		
Interpretation: Keep the null hypothesis		

Timed reading results

The average reading speed of the parsed texts among all 49 participants was 139.6 words per minute; 11.5wpm faster than the 128.1wpm taken by the same participants to read the block texts. These numbers are reported in seconds to completion in table 4 below. The data sounds impressive until one considers the fact that 36.7% (18) of all 49 participants actually read the block text faster; and, even more surprisingly, 54.2% (13) of the 24 participants from the PT group read the block text faster, while only 20% (5) of the 25 participants from the BT group read the block text faster. Furthermore, with statistics which read $t(48) = 1.600$, $p = 0.116$, these results suggest that no significant difference is being made between the reading rates across the two formats. As such, neither format has actually proven to lead to faster reading than the other.

Table 4: Statistical analysis on reading rates

$\alpha = 0.05$	Critical value: $t = 2.0106$ (two-tailed)	
	Block	Parsed
Average =	86.62 sec	79.50 sec
Std. Dev. =	5.52	28.95
Higher score distribution	5.52	< 31 (11)
All (PT only)		
	$t(48) = 1.600, p = 0.116$	
	Interpretation: Keep the null hypothesis	

Comprehension score results

Results from comprehension scores were equally baffling. Again, the BT group out-performed the PT group on the parsed reading questions. While 60% of the participants from the BT group scored the same or higher points for comprehension on the parsed text than on the block text, only 50% of the PT group could say the same. Although the scores were largely different, 81.6 points for the block text compared to 75.5 points for parsed, the two-tailed test, $t(48) = 1.597, p = 0.117$, once again implies that no significant difference was made between comprehension scores across the two formats. As such, neither format has actually proven to lead to better comprehension or reading efficiency than the other.

Table 5: Statistical analysis on comprehension scores

$\alpha = 0.05$	Critical value: $t = 2.0106$ (two-tailed)	
	Block	Parsed
Average =	81.63	75.51
Std. Dev. =	19.08	20.92
Higher score distribution	22 (12) > =14 (6)=	< 13 (6)
All (PT only)		
	$t(48) = 1.597, p = 0.117$	
	Interpretation: Keep the null hypothesis	

Eye tracking results

After data was collected, the eye movement patterns of 16 participants were compared in three sets of heat maps generated by the eye-tracking software (one set for scanning, one for skimming, and one for timed reading). Each set included four maps; one image for Group One looking at block text, another for Group One looking at parsed text, a third for Group Two looking at block text, and a fourth for Group Two looking at parsed text. Each heat map image was generated by a compilation of gaze traces from eight group members from each group whose eye movement patterns were properly recorded on all six pages that they were asked to scan, skim, or read. A proper recording was defined as one in which the eye tracking equipment maintained a constant and accurate trace for the same participant on every task that participant completed. Figure 3 below shows a proper gaze trace of a single 247

user which indicates exactly where he or she was looking over an elapsed time of just a few seconds. Other participants had plenty of proper gaze traces, but they were not consistent across every task due to excessive motion, excessive blinking, calibration problems, or other complications in data collection. Therefore, these 16 participants were the best choice.

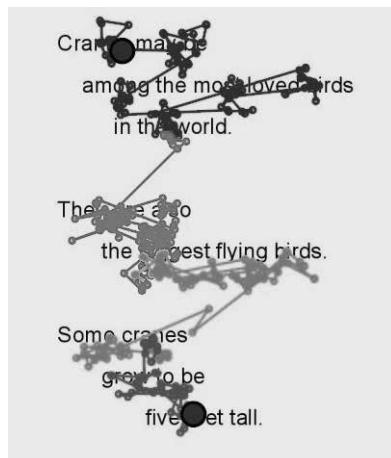


Figure 3. Gaze trace (single participant)

Scanning heat maps

When setting up the eye tracking equipment to record eye fixations within thresholds of 40 pixels and with a minimum duration of 15 milliseconds, it was easy to see where the participants looked and where they didn't during their rapid eye movements across the page for scanning, skimming, and timed reading. In Figures 4 and 5 below, two heat maps generated by the recordings of eight participants' eye movements during scanning exercises in two different groups are displayed. As each image is a compilation of eight participants' data points, the fixations which lasted for the shortest moments and those fixation points which did not fall in a region common to others do not show up here. However, in the original color images, degrees of intensity for each common fixation point are indicated by bright red for the longest and most common gaze fixation points, followed by orange, yellow, green, and blue, in that order. In the grayscale images pictured below, the outlying shades of gray are the areas that were colored blue or green, the white lines which appear to be outlining shapes were yellow, and the areas inside the white lines were shades of orange to red, where the darkest gray here indicates the brightest shade of red in the original image.

Cranes may be among the most loved birds in the world. They are also the biggest flying birds. Some cranes grow to be five feet tall. They wade in ponds and rivers.



Figure 5. Heat map of eye fixations while scanning syntactically parsed text.

In the eye movement recordings taken for the image in Figure 4, the average fixation duration was calculated to be 227.7ms. Therefore, it can be said that the words inside the white outlines were most likely looked at for more than a quarter of a second while the other words most likely were not. The heat maps in Figures 4 and 5 were recorded over the few seconds of time that the researcher had just called out the word “loved” and then “wade.” Since the participants had to find those two words, both of the words are hidden under the darkest shade of gray here. For the block text format, only two words come remotely close to warranting the same amount of attention, “biggest” and “flying,” which were buried under a shade of orange in the original image. However, for the parsed text format of the same section of text, these two words plus four more (also, cranes, five, and feet) draw the attention of the participants with fixation durations of more than the average 205.1ms for that recorded image.

Skimming data

The video screen captures of the participants’ eye movements as they skimmed the block text revealed a tendency for experienced skimmers to fixate on topic and closing sentences. However, the parsed text masked the topic sentences and encouraged the skimmer to quickly skim through as much as possible with fewer fixation points along the way. The number of fixation points while skimming for 15 seconds dropped from 26.9 for block text to 21.4 for parsed text on average for Group One and from 23.4 (block) to 16.3 (parsed) for Group Two.

Timed reading heat maps

The number of fixation points during timed readings also decreased noticeably. The number dropped from 134.5 for block text to 90.6 for parsed text on average for Group One and from 90.8 (block) to 76.3 (parsed) for Group Two. As seen by comparing Figures 6 and 7, fewer words are buried under the darker shades of fixation intensity in the parsed text than in the block text. This suggests that the readers have a tendency to observe larger language chunks in one fixation on parsed texts, while they take more time per individual word on block texts.

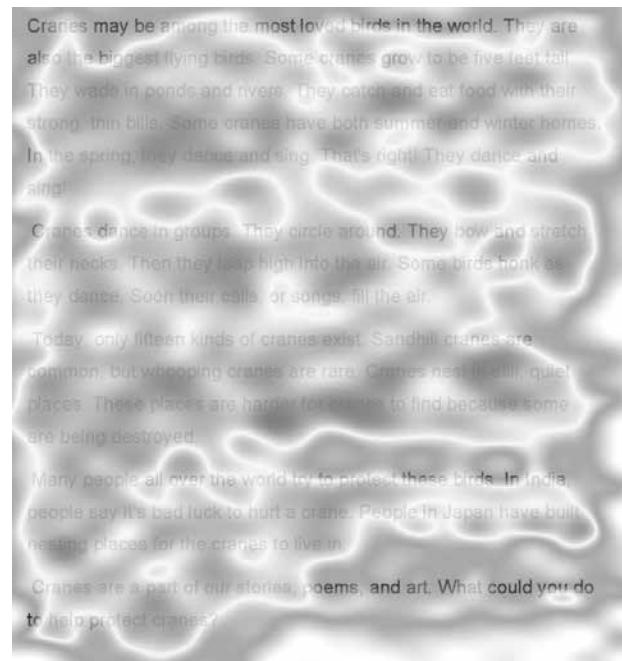


Figure 6. Heat map of eye fixations while reading block text.

Interpretation and limitations

Despite the challenges of acquiring a larger pool of reliable eye tracking records, the eye tracking data collected in this study adds much more to the interpretation of the results than a t-test or any other statistical analysis may offer. Below, the overall effects of using syntactically parsed text formats for scanning and skimming activities in an EFL context are further elaborated upon.

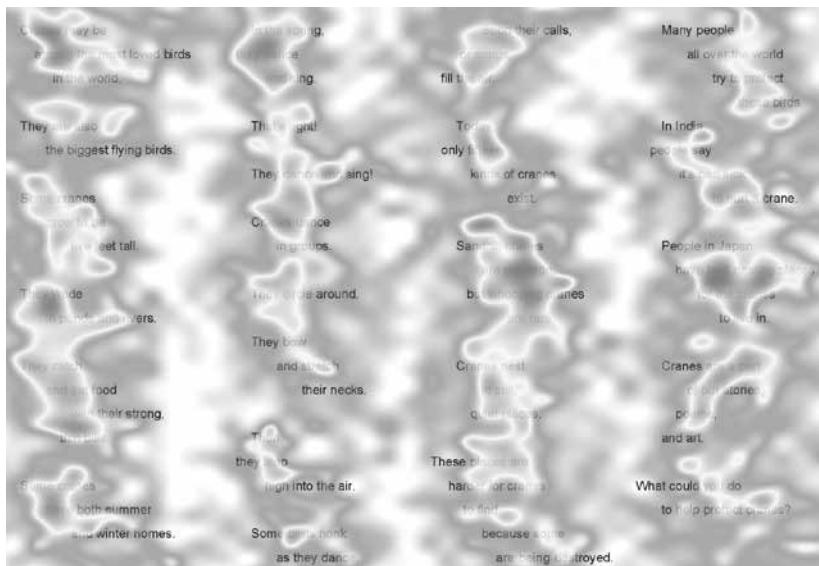


Figure 7. Heat map of eye fixations while reading parsed text.

More on the effects of parsed text on scanning and topic guessing

As seen in a comparison of Figures 4 and 5, the block text format lends itself to a wider range of words that are glanced over quickly while scanning than the parsed format. According to Samuels, Rasinski, and Hiebert (2014), a non-fluent reader needs 300ms to recognize a word and longer to comprehend it. As the white outline shapes in Figures 4 and 5 are estimated to represent fixations of about 200 to 250ms, and all fixations outside of the white outlines are estimated to have even shorter durations, the outlying words are assumed to have not been recognized at all. On the other hand, many of the much darker shaded words within the white outlines have fixation durations well over 300ms. Therefore, Figure 5 reveals a tendency, or at least a potential, for participants to recognize more words, irrespective of what they are looking for, during a parsed scanning activity than with a block scanning activity. This sheds light on the reason why the participants were able to guess the topic of the parsed text somewhat more quickly than the block text: Because they were actually taking in more vocabulary during scanning.

More on the effects of parsed text on skimming

Regarding the effects of syntactically parsed text on skimming, the participants were challenged by the necessity to skim a much wider plain of text rather than focusing on one intact set of words in a smaller field of view. Some participants complained about this and the fact that they could not find the topic and closing sentences within the parsed text. However, the parsed text also enabled the participant to glance at the text chunk-by-chunk. Therefore, fixating on one word would often equate to fixating on one chunk. While some 251

video screen captures of the parsed skimming activity suggested that the participants were glancing over most of the chunks in a reading, fixation points in related heat maps were mostly gathered around the first third of the text. At the same time, the video and heat map of the block skimming activity highlight the readers focusing on the first and last sentences of the first two paragraphs. The end results, regarding topic guessing after skimming, were fairly compatible between each format, but the pattern of eye movements to acquire the information was quite different.

Further discussion

When the 231 English conversation students in the six classes mentioned above finished all 20 intensive reading activities, each class was given one more activity in the format that they had not yet used for the activity. Then, they were asked in which format they preferred doing the activity. Among the 113 students in the PT group, only 21.2% (24) preferred block over parsed. However, 58.5% (69) of the 118 students in the BT group preferred block over parsed. This information alone shows that when the activity itself is presented in a nonconventional format, it takes time to become comfortable with the new format. As Warschauer (2011, June) has suggested, formats like VSTF do take some time for students to adapt to. Nonetheless, this study has stretched out its treatment period much longer than Yu and Miller (2010), Kanda (2012), and the author's two unpublished pilot studies. The study has also included a larger participant pool and a wider variety of activities. And, yet, the results are quite similar. There is still no conclusive evidence that a syntactically parsed text format can help classes of intensive EFL reading students in Asia to read English more efficiently. On the other hand, this begs the questions: Is syntactically parsed text more effective for some individuals than others? Does the ease of reading the syntactically parsed text actually slow down the more analytically minded technical college student, who might be having an "Ah-ha" moment in which they realize they can see the meaning more clearly in chunks; so, they initially slow down for taking it all in rather than moving forward and hoping it all makes sense as they continue along?

With respect to the first question, other researchers have had a similar inquiry. In a study focused on span limited tactile reinforcement (SLTR), a text format designed for helping dyslexic readers to read on tablets and smart phones, researchers found that left-justified parsed texts with only a few words per line help some readers to a significant degree but not others (Schneps, Thomson, Chen, Sonnert, & Pomplun, 2013). With this in mind, there is an apparent need for further research into VSTF and its derivatives. Does a syntactically parsed text help those with dyslexia? Is there a pattern that could noticeably immerge among different types of readers with different skill sets or different levels of reading disabilities?

Regarding the question about syntactically parsed text slowing down readers, occasionally, more regressive eye movements were witnessed in the video screen captures of the parsed text than the block text; but, the number of accurate recordings for determining the extent of these occurrences was too limiting to draw a conclusion here. Nevertheless, some of the same participants who were noted for these regressive eye movements and for reading the parsed text more slowly than the block were surprised to learn that they had performed better on the timed reading and questions for the block text activity than the parsed.

Concluding recommendations

Based on the observations made regarding scanning, if a reader is interested in capturing the gist of a reading or quickly obtaining a better understanding of how keywords may be used in context, a syntactically parsed text could help. However, for traditional instruction in skimming, in which developing English readers are taught to find key information by looking at topic sentences and closing statements, a parsed text format might be less desirable because it eliminates the ease at which such statements can be found in conventional paragraphs. Finally, since the effect of a parsed text format seems to differ between individual readers and contexts, and in answering to the voice of several research participants, it would be ideal if a reader could have a choice of format. On paper, that might appear to most as a wasteful project; but, in the form of an e-book reading, in which a push of a button would allow a reader to switch back and forth between formats, such an idea has merit.

Acknowledgements

The author would like to express his deep gratitude to Yoshitomo Matsubara and Gordon Bateson for their helpful technical support in the early stages of this project. Also, Yoshitomo Matsubara and Shogo Yasukawa deserve appreciation for their contribution to the initial reformatting of many pages of text for the project. Special thanks are extended to Mark Warschauer for his inspiration and personal advice on conducting research related to VSTF. And, finally, the author is grateful to the Ministry of Education, Culture, Sports, Science, and Technology in Japan for its financial support.

This work was supported by MEXT Grant-in-Aid for Scientific Research (C) Grant Number 24531166.

References

- Churchill, P. (2005). Who uses tools? In Don Parker (Ed.) *SRA Reading Lab 1b, Lime Power Builder 4*, McGraw-Hill Education.
- Demb, J. B., Boynton, G. B., & Heeger, D. J. (1997). Brain activity in the visual cortex predicts individual differences in reading performance. *Proceedings of the National Academy of Science*, 94, 13363–13366.
- Herbert, H., Herbert, J. C., & Bergman, C. L. (2003, August). Reading quickly, understanding clearly. *The Language Teacher*, 27 (8), 21–22.
- Kanda, A. (2012). EFL環境におけるVSTFによるリーディング活動についての効果と課題. Paper presented at the PC Conference, 2012, Kyoto. 195–198. Retrieved from: <http://gakkai.univcoop.or.jp/pcc/2012/papers/pdf/pcc028.pdf>
- Mirametrix, Inc. (2014). S2 eye tracker technical specifications. Last retrieved on July 31, 2014 from: <http://www.mirametrix.com/products/eye-tracker/>
- Samuels, S. J., Rasinski, T. V., & Hiebert, E. H. (2014). Eye movements and reading: What teachers need to know. *TextProject Article Series*, February 2014. Retrieved from: <http://textproject.org/library/articles/eye-movements-and-reading-what-teachers-need-to-know/>

- Schneps, M. H., Thomson, J. M., Chen, C., Sonnert, G., & Pomplun, M. (2013) E-Readers Are More Effective than Paper for Some with Dyslexia. *PLoS ONE* 8(9): e75634. doi:10.1371/journal.pone.0075634. Retrieved from: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0075634>
- Vendetti, A. (2005). *Dancers in the wild*. In Don Parker (Ed.) SRA Reading Lab 1b, Lime Power Builder 6, McGraw-Hill Education
- Walker , R. C., Gordon, A. S., Schloss, P., Fletcher, C. R., Vogel, C., & Walker, S. (2007). Visual-syntactic text formatting: Theoretical basis and empirical evidence for impact on human reading. Paper presented at the IEEE Professional Communication Conference, 2007, Seattle. 1-14. Retrieved from: <http://ieeexplore.ieee.org>
- Walker Reading Technologies, Inc. (2014). WebClipRead. Last retrieved on July 31, 2014 from: <http://www.liveink.com/>
- Walker, R., & Vogel, C. (2005). Live Ink®: Brain-based text formatting raises standardized test scores. Paper presented at the National Educational Computing Conference, 2005, Philadelphia. 1-39. Retrieved from: <http://www.liveink.com/>
- Walker, S., Schloss, P., Fletcher, C. R., Vogel, C. A., & Walker, R. C. (2005). Visual-syntactic text formatting: A new method to enhance online reading. *Reading Online*, 8(6). Retrieved from: <http://www.readingonline.org/>
- Warschauer, M. (2011, June). Re-imagining reading in digital learning environments. Plenary address to the JALTCALL annual conference, Kurume, Japan.
- Warschauer, M., Park, Y., & Walker, R. (2011). Transforming digital reading with visual-syntactic text formatting. *JALT CALL Journal* 7(3), 255-270.
- Yu, C.H., & Miller, R. C. (2010). Enhancing web page readability for non-native readers. Paper presented at the CHI Conference on Human Factors in Computing Systems, 2010, Atlanta. 1-10. Retrieved from: <http://groups.csail.mit.edu/uid/projects/froggy/chi10-froggy.pdf>

Author biodata

John C. Herbert is an Associate Professor of English at the National Institute of Technology, Akashi College. He is an active member of JALTCALL and the Moodle Association of Japan, where he promotes the use and development of open-source courseware for empowering teachers and students to gain better educational outcomes.