



Optimizing the reading of electronic text using rapid serial visual presentation

MONICA S. CASTELHANO and PAUL MUTER

Psychology Department, University of Toronto, Toronto, Ontario, M5S 3G3, Canada,
e-mail: muter@psych.utoronto.ca

Abstract. The focus on communications technology in recent years has led to the question of how to best display electronic text onto small-screened devices. Past studies have shown that the compact method of rapid serial visual presentation (RSVP) is efficient but not well liked. Two experiments were conducted to explore ways of improving the preference for and feasibility of RSVP. In experiment 1, the effects of a completion meter, punctuation pauses, and variable word duration were studied. Although the sentence-by-sentence and normal page formats were still superior, post-experiment ratings indicated that punctuation pauses improved user preference for RSVP, and its preference increased in general with practice. In experiment 2, a modified RSVP condition included a completion meter, punctuation pauses, interruption pauses and pauses at clause boundaries. This condition was significantly preferred to a normal RSVP condition. The present enhancements may increase the feasibility of using RSVP with small displays.

1. Introduction

Communications technology has exploded in the last ten years, bringing with it the problem of developing an optimal method of displaying electronic text on small screens. These screens are found on a variety of devices (e.g. cellular phones, pagers and desktop phones with screens). Many alternative formats for reading have been proposed, including: a moving window display, in which participants hit a button to produce the next words of the text, which is formatted normally otherwise (Just *et al.* 1982); times square, which is horizontal right-to-left scrolling of text (Chen *et al.* 1988, Kang and Muter 1989, Juola *et al.* 1995); line-stepping, which is similar to times square, but includes pausing at various intervals along the sentence (Bouma and de Voogd 1974); sentence-by-sentence presentation, in which the text is divided into sentences individually displayed (Moore and Zabrocky 1995, Rahman and Muter 1999);

and rapid serial visual presentation or RSVP (Forster 1970, Aaronson and Scarborough 1977, Potter 1984). RSVP involves displaying successive words of a text at a fixed point on the screen, usually at a predetermined rate.

RSVP was introduced by Forster (1970) to study the comprehension and processing of written language. It was later applied to computer screens (Aaronson and Scarborough 1977) and thought to be potentially superior to static displays (Juola *et al.* 1982, Masson 1983, Chen 1986). The advantage of RSVP was presumed to be the elimination of eye movements, which would lead to a possible reduction in cognitive load (Potter 1984).

Despite being a format with a wide range of potential applications (see Muter 1996), RSVP is extremely disliked in comparison to times square (Kang and Muter 1989), sentence-by-sentence and normal page format (Rahman and Muter 1999). The objective of the current paper was to investigate several methods of reducing the difficulties experienced during the reading process in RSVP. Six features were added to RSVP in various conditions: sentence-oriented processing; a completion meter; shorter presentation duration for common words; punctuation pauses; proposition integration; and interruption pauses.

1.1. Sentence-oriented processing

Most RSVP studies have acknowledged the importance of between-sentence pauses. However, most have merely added a 200 to 500 ms blank window in between sentence presentations (Masson 1983, Chen 1986, Juola 1988). There is evidence that in normal page reading, a longer pause is observed at the end of the sentence. Stine

(1990) compared processing in older and younger readers and discussed the significance of pauses at sentence boundaries. Sentence boundaries were shown to be especially important for younger readers, in that the most frequent and the longest pauses occurred there. Other studies exploring the reading process have also pointed to the correlation between time spent at the end of the sentence and the presence of factors such as infrequent words, the complexity of sentence structure and the integration of previously read information (Just and Carpenter 1980). Just and Carpenter assumed that post-processing of the sentence and integration of ideas takes place at sentence boundaries.

As seen in the studies above, the amount of time a reader spends at the end of a sentence can vary greatly. Instead of relying on a predetermined end-of-sentence pause, the present study used a more interactive approach. That is, as in Rahman and Muter (1999), one sentence was presented at a time, and a key press prompted the presentation of the next sentence, thus allowing for variation in pauses and an increase in user control. In addition, subjects were permitted to repeat the current sentence.

1.2. Completion meter

Peripheral cues allow a reader to judge what has been read and the amount of reading remaining to be done. Rothkopf (1971) found incidental memory for location within a page may facilitate processing during normal page reading. This cue is absent in RSVP and the feeling of being 'lost' may contribute to the uneasiness felt by users. O'Hara and Sellen (1997) investigated different methods of increasing location memory by manipulating how text was read from a small screen. They found that using a thumbnail of the entire passage that highlighted the current sentence being read increased the likelihood that the location and information would be later recalled. With similar aims, Rahman and Muter (1999) implemented a completion meter and found user preference of RSVP increased. The current experiment added a completion meter in some RSVP conditions. The completion meter is a graphical representation of the whole passage that changes as each sentence is read. The meter represents the unread sentences, including the present sentence, as a vertical bar (|), which changes to a dot (.) after the sentence is read (see figure 1C).

1.3. Common word duration

Most studies of RSVP present all words for a constant duration, but reading studies have identified variable

A

Visual recognition involves storing and retrieving memories. Neural activity, triggered by the eye, forms an image in the brain's memory system that constitutes an internal representation of the viewed object. When an object is encountered again, it is matched with its internal representation and thereby recognized. Controversy surrounds the question of whether recognition is a parallel, one-step process or a serial, step-by-step one. Psychologists of the Gestalt school maintained that objects are recognized as wholes in a parallel procedure: the internal representation is matched in a single operation. Other psychologists have proposed that internal representation features are matched serially with an object's features. Although some experiments show that, as an object becomes familiar, its internal representation become more holistic and the recognition process correspondingly more parallel, the weight of evidence seems to support the serial hypothesis, at least for objects that are not notably simple and familiar.

B

Psychologists of the Gestalt school maintained that objects are recognized as wholes in a parallel procedure: the internal representation is matched in a single operation.

C

....|||
representation

Figure 1. Panel A: the normal page format with the full text of one of the passages; Panel B: the sentence-by-sentence showing a portion of the same passage used in figure 1A; Panel C: An excerpt of the sentence in figure 1B shown in the RSVP format with a completion meter. The completion meter shows that the participant is currently on the fifth sentence of seven. The bars represent the sentences to be read in the passage, including the present sentence, while the periods reflect those that have been read.

processing times as a function of word type. Just *et al.* (1982) reported that although content words were fixated 83% of the time, function words were fixated only 35% of the time in normal page reading. The processing of function words may occur in their shape, not in the serial interpretation of each letter, because of their high frequency in the English language (Just and Carpenter 1980). Proofreading and skimming studies also point to a more global processing of the function words (Reicher 1962, Healy 1976, Haber and Schindler 1981). For the common word duration modification, in the first experiment, the duration of 11 high frequency words was reduced in some conditions.

1.4. Punctuation pauses

The significance of punctuation pauses (other than for periods) in RSVP studies has been largely overlooked, but reading research has pointed to the importance of

pausing for the post-processing and integration of information (Just and Carpenter 1980). Stine (1990) found that younger and older groups spent a greater amount of time at both major and minor clause boundaries. A study on oral reading has shown that 91% of pauses occur at punctuation points (O'Connell and Kowai 1986). By providing a slightly longer duration for words with punctuation in RSVP, extra processing can occur during the sentence presentation instead of at the end of the sentence.

1.5. Proposition integration

Another way of potentially improving RSVP is through emphasis on the propositional meaning in a sentence. The relationship between the propositions of a text is often cued by words such as 'because' (which indicates an integration) or 'although' (which indicates awaiting more information before the integration is made). A listener or reader emphasizes these words by spending more time processing them (Townsend 1983). In the psycholinguistic literature, the importance of such words in reading and listening has been shown as: segmentation devices, in that the word indicates the place where the reader has to start a new syntactic structure (Clark and Clark 1977, Noordman and Vonk 1997); integration devices, which indicate that the current information has to be integrated with previous information, as mentioned above (Townsend 1983, Townsend *et al.* 1987); and as inference devices, which, according to Noordman and Vonk (1997), trigger a process that derives a general causal relationship between events in the different clauses. By allotting extra processing time for these words, RSVP readers would be better able to integrate information at clause boundaries.

1.6. Interruption pauses

The ability to pause word presentation in mid-sentence is a privilege that RSVP readers have not had. By allowing the participants to stop at any point in time during the sentence presentation, they may feel a greater sense of control and less frustration. Additionally, when a word near the beginning or middle of a sentence is misunderstood or unfamiliar, rereading the sentence would be more helpful at that point in time as opposed to waiting until the end of the sentence. Therefore, when the sentence presentation was paused, subjects were given the option of regressing to the beginning of the sentence. Muter *et al.* (1988) found regressions to the beginning of the

sentence to be more frequently used than 8-, 4-, or 2-word regressions.

2. Experiment 1

In experiment 1, the RSVP conditions had three possible modifications as discussed above: (1) a completion meter was added; (2) common word duration ('I', 'he', 'she', 'it', 'you', 'the', 'a', 'to', 'of', 'and', 'in' according to Taylor 1976) was shortened to half the presentation time of other words (normal words were presented for 230 ms, while common words were presented for 115 ms); and (3) punctuation pauses were added to present words with punctuation (,; \ = / * ' () - &) for twice as long as other words. All possible combinations of the three modifications were tested for a total of eight RSVP conditions. For comparison purposes, a normal page condition and a sentence-by-sentence condition were included. Sentence-by-sentence has been minimally investigated, but is highly compact and shows promise in terms of efficiency, as measured by reading speed and comprehension (Rahman and Muter 1999).

2.1. Method

2.1.1. *Participants.* There were 20 participants, 12 females and eight males. Age ranged from 19 to 24 years, with a mean age of 20.4 years. All were undergraduates who completed the experiment for a credit in an introductory psychology course.

2.1.2. *Materials and apparatus.* Participants read 10 passages from the comprehension portion of the Graduate Record Examination study guide for the General Test (1996). The passages averaged 153.3 words and ranged from 136 to 173 words. They were displayed in black on a white background on a PowerPC Macintosh. The display was 17 inches (diagonally) with a resolution of 800 × 600 pixels. The text was presented in a 12-point serif font called 'Times'. The function keys were both located on the number pad, on the right side of the keyboard. The '*' key was labelled with a blue square sticker, while the '/' key, immediately to its left, was labelled with a red square sticker. Participants read the passages in three different formats: normal page, sentence-by-sentence and RSVP (see figure 1). The normal page format was single spaced and was 6.5 inches across. The sentence-by-sentence condition was similar except that each sentence was presented individually and started half way down the screen.

In the RSVP conditions words were centred on the screen both horizontally and vertically and were

displayed at a within-sentence rate of approximately 260 words per minute (WPM); all words were presented for 230 ms, unless they belonged to one of the conditions mentioned above.

2.1.3. *Design.* In all, there were ten conditions: normal page, sentence-by-sentence, RSVP without modifications, RSVP with completion meter only (CM), RSVP with shortened common word duration (WD) only, RSVP with punctuation pauses (PP) only, RSVP with CM and WD, RSVP with CM and PP, RSVP with PP and WD, and finally RSVP with CM, WD, and PP. All participants performed in every condition.

2.1.4. *Procedure.* The order of the ten conditions was randomized for each participant. Similarly, the order of the ten passages was randomized for each participant. At the end of each trial, the experimenter would set up the next condition, while the participant answered a mock comprehension question. It was included to keep participants focused on the passage content and from skipping over sentences for a hasty finish. Participants would then give a magnitude estimation (Stevens and Galanter 1957) as a preference rating for the format just seen. They were instructed to assign a number that appropriately conveyed the extent of their like or dislike of the condition. Each condition was to be compared to the liking of reading from a book page, designated as 100. For example, if the condition was liked twice as much as a book page, then the response would be 200. If the condition was liked only half as much, the response would be 50.

The experiment took approximately one hour to complete and each participant was tested in isolation. Participants were given a verbal description of all ten conditions; however, at the beginning of each trial, they were not informed which format they would be viewing. They were informed that their reading speed was being recorded and were told to read as quickly and accurately as possible. At the beginning of each trial, a message appeared in the centre of the screen in red letters that read, 'Press any key to start'. Once a key was pressed, the timer was activated and the trial commenced.

The normal page condition simply required the participants to read the passage presented in full on the screen and then to press any key when finished. Participants were instructed to do so during the initial instructions and in a written message at the bottom of the screen during this trial. The message read, 'Press any key when finished'. At the end of each trial, a message would appear in the centre of the screen reading 'Please call experimenter to set up the next trial', at which time the experimenter would return to set up the next condition.

In the sentence-by-sentence condition, the blue key allowed participants to advance to the next sentence. The red key did nothing and no regressions were permitted. The sentence would remain on the screen until the blue key was pressed, which meant that the current sentence could be read as often as the subject wished.

Similarly, in all the RSVP conditions the blue key would advance to the next sentence. Because words were displayed only one at a time, the option of rereading the sentence was offered, consistent with the other conditions. The red key served to repeat the current sentence, starting at its beginning. Participants could read the same sentence without limitations on how often it was repeated. Regressions to previously read sentences were not permitted in any RSVP condition. The last word of each sentence remained on the screen until one of the function keys was pressed. Words were presented within sentences at speeds of approximately 260 wpm, with the exception of the conditions with punctuation pauses, which increased the presentation time, or with common word duration, which decreased the presentation time of the eleven chosen high-frequency words.

In figure 1, the screen displays for each of the three formats are shown. Figure 1A shows the normal page condition for one of the passages used in the experiment. A portion of the same passage was used in figure 1B in the sentence-by-sentence format. An excerpt from the sentence is shown in the RSVP condition with a completion meter in figure 1C. The completion meter indicates that the subject is presently on the fifth sentence of seven. The bars represent the sentences to be read in the passage, including the present sentence, while the periods reflect those that have been read. Every time the blue key is pressed to advance to the next sentence, the bar of the current sentence is replaced by a period.

In all conditions the final key press stopped the timer and ended the trial with the final screen and message described above. Reading speed was calculated in terms of words per minute (wpm).

In addition to the magnitude estimation ratings given at the end of each trial, at the end of all ten trials, participants gave post-experiment ratings. They were asked to rate the three basic modifications of RSVP according to a 9-point scale, where 1 indicated extreme dislike and 9 represented extreme liking.

2.2. Results

The analyses adopted the recommendations of Stevens (1996: 448–454) to control for possible violations of the sphericity assumption for within-subject designs. For all comparisons, when the Greenhouse–

Geisser and the Huynh–Feldt epsilon estimate averaged above 0.70, Tukey post hoc tests were used; when the estimate was less than 0.70, Bonferroni tests were used. In either case, the overall alpha was 0.05; therefore, the alpha for each individual test was much less.

2.2.1. Preference ratings for each of the ten conditions. Figure 2 shows the means and standard errors of the preference ratings. A one-way completely within-subjects analysis of variance (ANOVA) was performed on the preference ratings, conveying a significant difference among the ten format conditions, $F(9, 171) = 8.47$, $p < 0.01$ (with the Greenhouse–Geisser epsilon correction, $p < 0.01$). The Bonferroni tests revealed a significant difference between the normal page and all the RSVP conditions, in that the normal page was preferred more than any of the RSVP conditions. The sentence-by-sentence condition was also found to be significantly more liked than all RSVP conditions, but there was no significant difference between it and the normal page condition. None of the RSVP conditions differed significantly from each other.

2.2.2. Preference ratings by trial order. The means for each of the ten trials were calculated, including only the eight RSVP conditions, in order to look for practice effects. Because the conditions were in a randomized order, there were not an equal number of observations for each trial. In figure 3, the means and standard errors are shown. A within-subjects ANOVA revealed a significant effect for trial order ($F(9,131) = 3.90$,

$p < 0.01$): Ratings increased as subjects spent more time reading in the RSVP format. Further analysis showed a significant linear regression ($F(2, 157) = 10.78$, $p < 0.01$; slope = 4.32; intercept = 31.37).

2.2.3. Post-experiment ratings. The post-experiment ratings for the each of the three basic RSVP modifications were also analysed. Figure 4 shows the means and standard errors of the three modifications. There was a significant effect, $F(2, 38) = 6.84$, $p < 0.01$ (Greenhouse–Geisser epsilon correction, $p < 0.01$). Tukey tests indicated that punctuation pauses (PP) differed significantly from both common word duration (WD) and completion meter (CM). However, WD and CM did not differ significantly. Punctuation pauses were liked, but the other modifications were given neutral scores.

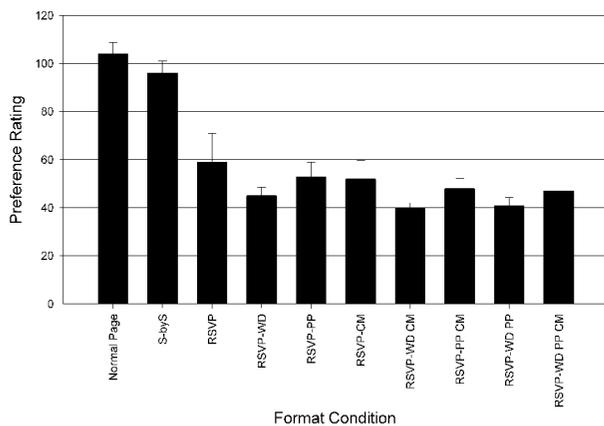


Figure 2. Means and standard errors for the preference ratings of the ten conditions of experiment 1. The ten conditions were normal page, sentence-by-sentence RSVP without modifications, RSVP with completion meter only (CM), RSVP with shortened common word duration (WD) only, RSVP with punctuation pauses (PP) only, RSVP with CM and WD, RSVP with CM and PP, RSVP with PP and WD, and finally RSVP with CM, WD, and PP.

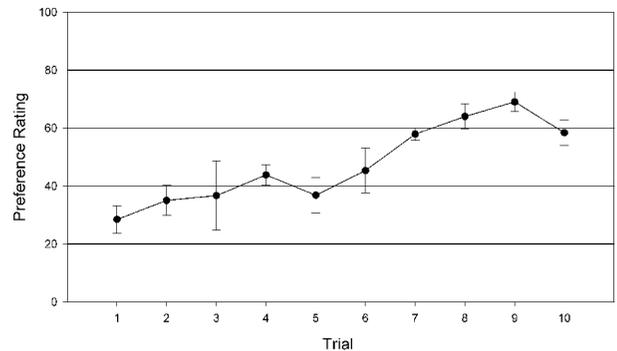


Figure 3. The means and standard errors for RSVP preference ratings according to trial order for experiment 1.

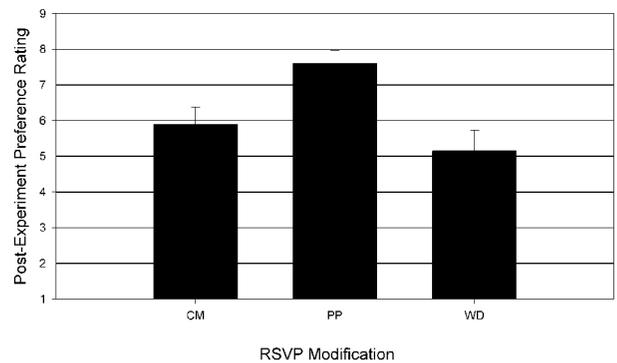


Figure 4. Means of the post-experiment ratings for the three basic RSVP modifications with standard errors for experiment 1. The three modifications were: completion meter (CM), shortened common word duration (WD) and punctuation pauses (PP). Post-experiment ratings were according to a 9-point scale, where 1 indicated extreme dislike and 9 represented extreme liking.

2.2.4. *Reading speed.* A within-subjects ANOVA was run on the reading speeds, which were calculated as words per minute (WPM) for each of the formats. There were no significant differences among the conditions ($F(9, 171) = 0.26, p > 0.1$).

2.3. Discussion

Although no effects of the individual RSVP modifications were observed in the immediate post-trial judgments, the post-experiment ratings of the individual RSVP modifications show that punctuation pauses are valued. In addition, over just ten trials, preference for RSVP gradually increased with practice. The results show that the space-saving sentence-by-sentence format was as well liked as the normal page format.

Because of the results of past studies (Juola 1988, Muter *et al.* 1988), the pace of word presentation within-sentence was set at 260 WPM. In theory, the advantage of RSVP is in its elimination of the saccades necessary in normal reading. Efficiency scores may have decreased in past studies with slower rates because of the unnecessary movement of the eyes during the presentation of each word. The problem with setting the rate at a fixed high speed is the disregard for individual differences. A person who naturally reads more slowly than 260 WPM would have trouble understanding and keeping up with the sentences being presented. In experiment 1, one of the modifications (WD) sped up the presentation of some words, which might have in effect aggravated the problem. (In addition, the preference for punctuation pauses may have been because it reduced the within-sentence presentation rate, and so allotted extra processing time.) Alternatively, the presentation duration of each word varied according to the condition and according to word type; adding the WD condition and the PP condition may have decreased the predictability of word presentation, making reading from RSVP more uncomfortable than it would be otherwise.

3. Experiment 2

The second experiment likens RSVP to listening to speech, and so introduces manipulations that emphasize the rhythm of word presentation. Although there are opposing views as to how closely related reading and listening are, many researchers have demonstrated that at the comprehension level, they are quite different. Speech is considered to have much richer prosodic cues than written text (Fries 1962, Goldman-Eisler 1972). In addition, the listener can ask the speaker to stop at any time. These factors can now be incorporated into

reading because computers allow for dynamic text presentations. The words still require decoding, but the presentation may be varied in tempo, allowing for emphasis to be placed on certain words. The present experiment attempts to improve RSVP preference ratings by using an approach to reading that is essentially a hybrid of listening to speech and reading.

In experiment 2, one condition using the RSVP format (modified-RSVP) differed in three ways: punctuation pauses as in experiment 1 (punctuated words presented for 460 ms); interruption pauses (pausing at any point during sentence presentation to allow for increased processing time); and proposition integration (presentation time of 460 ms for words such as 'if', 'because', 'after' and 'although' that predict and integrate propositions). These three alterations were combined into a single condition, in a manner similar to that used by Muter and Maurutto (1991, experiment 2). This 'kitchen sink' approach does not permit the contributions of each modification, whether positive or negative, to be determined. However, it has advantages over experiment 1 in that it increases power, while allowing any positive interactions between the modifications to take place. The modified-RSVP condition was compared to plain-RSVP (standard version) and to a normal page format.

3.1. Method

3.1.1. *Participants.* A total of 18 undergraduate students participated in the second experiment. There were seven males and 11 females. The ages ranged from 19 to 36, with a mean of 25.4 years. All completed the experiment for a credit toward an introductory psychology course.

3.1.2. *Materials and apparatus.* These were identical to those used in experiment 1, except that in the modified-RSVP an extra function key was needed to enable the pause/un-pause option. The functions were located on the top-right corner of the number pad. The '=' key was labelled with an arrow pointing to the left, the '/' key was labelled with an arrow pointing to the right, the 'P' key was labelled with a 'P' and the '*' key was labelled with an arrow pointing to the right.

All passages were parsed manually. Pauses were placed at the beginning of new clauses. For example in the sentence: 'It would be wonderful to observe a singularity and obtain direct evidence of the undoubtedly bizarre phenomena that occur near one', a pause was placed at the word 'singularity' and 'phenomena' to signal the beginning of the second and the third proposition, respectively. These selected words were presented for twice the duration of normal words, unless

they contained a punctuation mark, in which case a pause has already been added to that word. Parsing text via computer programs can be done and has already been achieved to a modest degree in the past (Granaas 1985, Postigo *et al.* 1992). Based on the speed with which new technological advances take place, an improved method is a reasonable aspiration for the near future.

The comprehension test (see below) was composed of three multiple-choice questions for each passage. These questions were mostly inferential, with some answers requiring specific facts from the passage.

3.1.3. *Design.* There were three conditions: normal page, plain-RSVP (without modifications), and modified-RSVP with punctuation pauses, propositional integration, and interruption pauses. All participants completed every condition.

The experiment took place in two sessions: the practice session (in which the subject practised reading from the plain-RSVP for ten trials) and the test session (in which the three conditions were presented in blocks of five trials each). Only the plain-RSVP condition was used in the practice session in order to give all participants a base from which they would be able to judge the modified-RSVP condition. This way, any preference for modified-RSVP in the second session could not be attributable to familiarity.

3.1.4. *Procedure.* In the practice session, subjects read 10 randomly arranged passages in the plain-RSVP format. The passages varied from 139 words to 164 words, with an average length of 152.4 words. At the end of each trial the experimenter would set up the next condition, while the participant answered three multiple-choice comprehension questions. Participants would then give a magnitude estimation, according to the same instructions as in experiment 1. The practice session took about 40–60 minutes to complete.

In the test session, subjects read five randomized passages in each of the three format conditions; thus, 15 passages were read in blocks of five. The passages had an average length of 155.1 words and ranged from 136 words to 173 words. The order of the three conditions was counterbalanced across subjects; all six possible orders were used three times each. Participants were given a verbal description of all three conditions at the beginning of the session. They were informed that their reading speed was being recorded and were told to read as quickly and accurately as possible. The testing session was about one hour to one hour and 15 minutes.

In both sessions, the display did not differ from experiment 1, with the exception of a few details. The plain-RSVP condition was similar to the presentation

format used in the first experiment, except that a completion meter was added (see figure 1C).

The buttons used to control the display of the sentences differed from the first experiment. That is, the 'left arrow' key was used to repeat the current sentence during the presentation of either RSVP condition, and the 'right arrow' key was used to move to the next sentence. The 'P' key functioned only in the modified-RSVP condition and was used to stop the presentation of the sentence at any point in time. Subjects could resume the presentation of the sentence in the same location by pressing the same key or could restart the presentation of the current sentence by pressing the 'left arrow' key.

The dependent variables were preference rating and efficiency score. The efficiency score was calculated by multiplying the reading speed (WPM) by comprehension score (percentage correct).

3.2. Results

3.2.1. *Test session.* A one-way repeated measures ANOVA performed on the preference ratings for the three conditions revealed a significant difference among

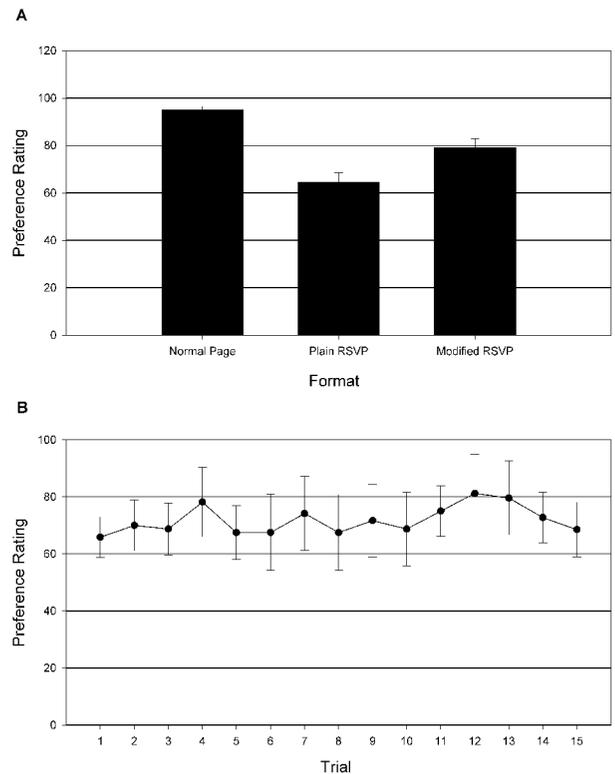


Figure 5. Means and standard errors for the Test Session in experiment 2: (A) the preference ratings by format condition; (B) the preference ratings by trial order.

them ($F(2, 34) = 8.81, p < 0.01$; with Greenhouse–Geisser epsilon correction, $p < 0.01$). The Tukey tests showed that all three conditions were significantly different from one another. That is, as shown in figure 5A, the normal page condition was significantly more liked than the modified-RSVP, which in turn was significantly more liked than the plain-RSVP condition. No practice effect was observed for preference ratings over the 15 trials of the test session in the RSVP conditions (see figure 5B): ($F(2, 238) = 0.54, p > 0.1$).

Figure 6 shows the efficiency score (reading speed \times comprehension) means for each of the formats. A within-subjects ANOVA indicated no significant difference among the three conditions ($F(2, 34) = 2.52, p > 0.1$). Reading speed, one component of the efficiency score, showed a significant difference within the format conditions ($F(2, 250) = 26.36, p < 0.01$). A Tukey test revealed that although the normal page format did not differ from the plain-RSVP condition, the modified-RSVP was significantly slower than both. Figure 7A shows the mean reading speed for each format in terms of words per minute (WPM).

The comprehension score was calculated from the percentage correct on the multiple choice questions. Figure 7B shows the mean comprehension scores for each format in terms of percentage correct. Despite the fact that the modified-RSVP seems to have a higher score overall, a one-way repeated ANOVA did not show a significant difference within the conditions ($F(2, 250) = 2.53, p > 0.05$).

3.2.2. *Practice session.* Figure 8 shows the mean preference ratings for the plain-RSVP for the ten practice trials. A within-subjects ANOVA revealed a significant difference among the trials ($F(9, 153) = 2.12, p < 0.05$). Although there seems to be some variance in how the format was rated through the practice trials, a

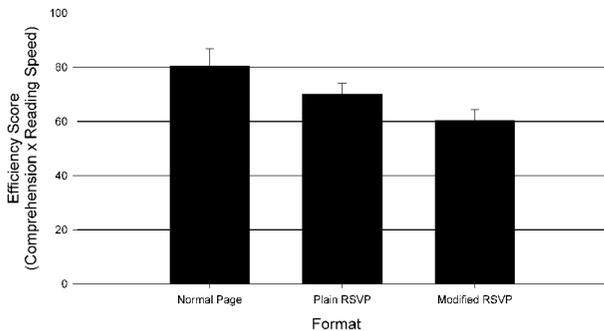


Figure 6. The efficiency (reading speed \times comprehension) means and standard errors by format for the testing session in experiment 2.

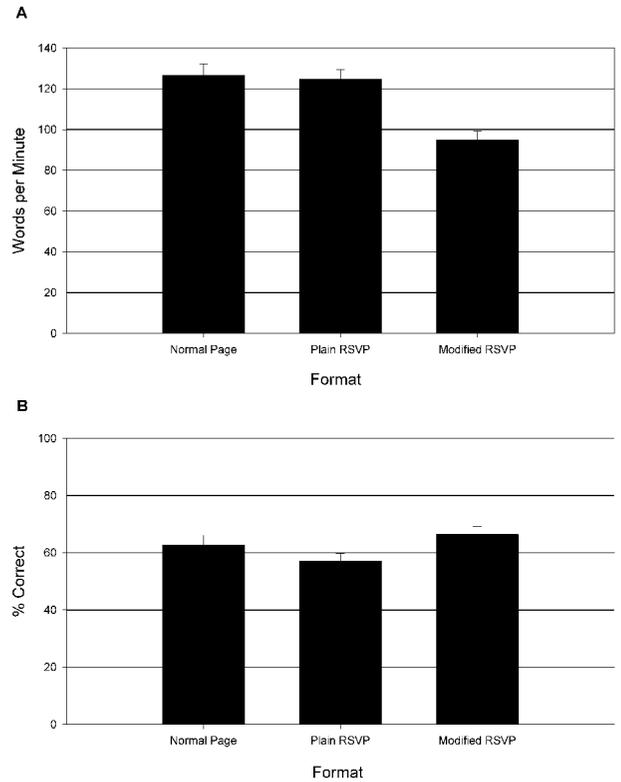


Figure 7. Efficiency scores for each format are broken down into its components: (A) reading speed (words per minute) by format condition; (B) comprehension scores (percentage correct) by format condition.

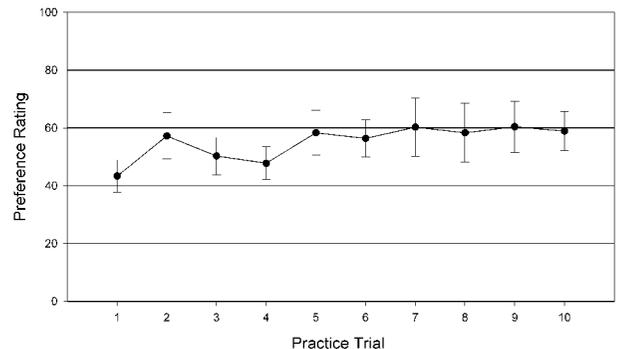


Figure 8. The preference ratings means and standard error for the plain-RSVP in the ten practice trials.

linear regression did not reveal any evidence of a practice effect ($F(1,178) = 1.243, p > 0.1$).

The efficiency scores for the practice sessions revealed the same pattern of results as the preference ratings. Despite a significant overall F-test ($F(9,153) = 2.63, p < 0.01$), a linear regression analysis did not show any evidence of a practice effect ($F(1,178) = 2.947, p > 0.05$).

Tukey tests revealed that the trial 10 and 4 were significantly higher than trial 3.

3.3. Discussion

In the second experiment, the modified-RSVP improved the preference ratings as compared to plain-RSVP. Because the modified-RSVP incorporated a number of changes, it is not possible to determine which modification or modifications contributed to this improvement. However, as discussed above, the design of experiment 2 allowed an increased chance of improving RSVP. Any improvements, regardless of whether they take place from the collection of enhancements as a whole or from a specific one, indicate that RSVP can be improved and made more feasible.

The practice session consisted of plain-RSVP only. The preference ratings for the modified-RSVP condition in the second session may have been even higher if subjects had also experienced a session of practice with modified-RSVP.

In the practice session, the practice effects seen in experiment 1 did not occur in either the preference ratings or in the efficiency scores. The reason for this is not known, but one can speculate that the demand characteristics in the two experiments were different. That is, although subjects were reading in the RSVP format for eight of the ten trials in the first experiment, they were perceived as being different from each other. However, during the practice session in the second experiment all passages were read in the exact same RSVP format. The perceived differences among the RSVP formats in experiment 1 could have given subjects ample reason to see the later-viewed passages as somehow better than the first ones, whereas in experiment 2, the knowledge that all formats were identical may have influenced subjects to rate the format consistently across the trials.

4. General discussion

The present experiments suggest that although the standard RSVP format is disliked, certain modifications do improve preference rating for RSVP, even with very little practice. This enhancement can be seen in the practice effect of the first experiment, in which ratings tended to increase as the number of trials completed increased. Also, in the first experiment, the post-experiment ratings of the modifications used in the RSVP formats showed a significant preference for the punctuation pauses. The concept of allotting more processing time at certain points of the sentence

presentation was taken a step further in the second experiment. Results demonstrated a significant increase in the preference rating of the modified-RSVP over the plain-RSVP format.

The improvement in RSVP preference in experiment 2 is attributable to the combination of revisions chosen for the modified-RSVP condition that imitated speech characteristics. The modified-RSVP incorporated three modifications: punctuation pauses, proposition integration and interruption pauses. Evidence that these changes may improve reader's comprehension can be seen in the psycholinguistic literature. A study by Calvo *et al.* (1999) manipulated processing times for explicit comprehension and inference making by using a RSVP paradigm and varying the time allotted for processing while the sentence was being read and the time allotted between sentences. A priming test was used to detect the activation of explicitly presented information and implied information. Calvo *et al.* found that decreasing within-sentence presentation rate tended to improve online processing of explicitly presented information, but it had no effect on the production of inferences. On the other hand, increasing the delay between sentences tended to increase the likelihood that an inference was drawn, while it did not affect the representation of the explicitly represented information. These results imply that the sentence-oriented approach combined with the delays associated with propositions and punctuation may help to improve the comprehension of material, both explicit and implicit. Further improvements of the RSVP format may be found in a combination of enhancements supported by the reading literature and the psycholinguistic literature.

As described above, many RSVP paradigms have not been able to successfully integrate increased control for the user. Attempts to allow users to control presentation rate have proven to be disastrous in terms of efficiency (Juola 1988, Muter *et al.* 1988). Muter *et al.* attempted to allow for increased control by permitting sentence regressions and self-pacing. Subjects were allowed to regress two, four or eight words or to the beginning of the present sentence and controlled the speed of within-sentence word presentation. Results showed a significant decrease in reading speed, without compensation in the form of improved comprehension scores. As an alternative to this type of modification, the present study included a pause button. In addition to the improvements in comprehension that the extra processing time allows, there is a greater amount of control given to the user. It is possible that this factor alone may have lead to the improvement in preference rating of the modified-RSVP over the plain-RSVP condition.

The comparison between the normal page and the modified-RSVP condition shows some room for future improvements. The following are some possible limitations of the present study that may be explored in future research.

One reason the preference ratings may have not improved further was the variation in individual reading rates. Past studies have suggested that decreasing speeds, to cater to individual differences, may remove the theoretical advantage of eliminating saccades (Muter *et al.* 1988, Rahman and Muter 1999). Therefore, the within-sentence presentation rate was kept artificially high. Future research can be aimed at improving allowance for individual differences without losing the theoretical advantage of RSVP.

Another factor may be the insufficient time to gain familiarity with the format and improve the participants' level of competence. The skills required to read in RSVP are unfamiliar to users. In the first experiment, the greater liking of sentence-by-sentence presentation could be explained because it uses many of the same skills as are used in normal page.

A problem may lie in the theoretical assumption that RSVP is advantageous because it reduces cognitive load by eliminating eye movements. To date, there is little evidence to support this assumption. Rubin and Turano (1992) examined the cost of saccadic eye movements on reading time by comparing RSVP to normal page reading. They found that RSVP reading rates were consistently higher, with adequate comprehension levels. However, few participants felt comfortable, and reading at high speeds using RSVP was 'taxing' for most participants (1992: 901). Therefore, even if RSVP does not necessitate the programming and execution of saccades, the concentration levels needed may, in fact, exceed the effort required for eye movements in normal reading. By requiring eye movements to be suppressed, RSVP may actually increase cognitive load and divert attention from the content of the passage being read (Bouma and deVoogd 1974).

Future studies may expand the enhancements made in the present study to other methods of presenting electronic text. One example is the times square format, in which text is continuously scrolled across the screen. Kang and Muter (1989) noted that this format was preferred to RSVP, but its efficiency was inferior. They improved its efficiency by simply modifying the smoothness of the scrolling. An alternative way of finding a format that is both efficient and likable may lie in improving a format that is already familiar and liked.

In addition to the feasibility of using RSVP as an alternative reading method for the normal reading population, its usefulness may be seen in other areas.

Fine *et al.* (1996) studied the possible advantages of RSVP on an aging reading population. They found that, unlike sentence-oriented displays, the benefits of RSVP did not decrease as visual acuity and contrast sensitivity decreased. Other special populations have also been proposed to benefit from RSVP displays, such as dyslexics (Potter 1984), and those who have *retinitis pigmentosa* or other peripheral vision impairments (Williamson *et al.* 1986). RSVP has also been shown to be potentially beneficial to less efficient readers (Chen 1986, Juola *et al.* 1987).

By including the above-mentioned enhancements, the widespread application of RSVP to electronic devices with small screens is possible. Additionally, research using the RSVP paradigm can address theoretical questions regarding the reading process and its potential applications in various populations.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada. The authors wish to thank Douglas Creelman and Tarjin Rahman for their assistance and generosity, and Henrietta Lempert, Tarjin Rahman, Fernanda Ferreira, Thomas Carr and John Henderson for comments on earlier drafts.

References

- AARONSON, D. and SCARBOROUGH, H. S., 1977, Performance theories for sentence coding: Some quantitative models. *Journal of Verbal Learning and Verbal Behaviour*, **16**, 277–303.
- BOUMA, H. and DEVOOGD, A. H., 1974, On the control of eye saccades in reading. *Vision Research*, **14**, 273–284.
- CALVO, M.G., CASTILLO, M.D. and ESTEVEZ, A., 1999, On-line predictive inferences in reading: Processing time during versus after the priming context. *Memory and Cognition*, **27**, 834–843.
- CHEN, H.-C., 1986, Effects of reading span and textual coherence on rapid-sequential reading. *Memory and Cognition*, **14**, 202–208.
- CHEN, H.-C., CHAN, K.-T. and TSOI, K. C., 1988, Reading self-paced moving text on a computer display. *Human Factors*, **30**, 285–291.
- CLARK, H. H. and CLARK, E. V., 1977, *Psychology and Language* (New York: Harcourt).
- FINE, E. M., PELI, E. and REEVES, A., 1996, Simulated cataract does not reduce the benefit of RSVP. *Vision Research*, **37**, 2639–2647.
- FORSTER, K. L., 1970, Visual perception of rapidly presented text segments of varying complexity. *Perception and Psychophysics*, **8**, 215–221.

- FRIES, C. C., 1962, *Linguistics and Reading* (New York: Holt).
- GOLDMAN-EISLER, F., 1972, Pauses, clauses, and sentences. *Language and Speech*, **15**, 103–113.
- GRANAAS, M. M., 1985, Simple, applied text parsing. *Behavior Research Methods, Instruments and Computers*, **17**, 209–216.
- GRE, 1996, *GRE: Practising to Take the General Test*, 9th edn (Princeton, NJ: Educational Testing Service).
- HABER, R. N. and SCHINDLER, R. M., 1981, Error in proof-reading evidence of syntactic control of letter processing? *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 573–579.
- HEALY, A. F., 1976, Detection errors on the word The: evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 235–242.
- JUOLA, J. F., 1988, The use of computer displays to improve reading comprehension. *Applied Cognitive Psychology*, **3**, 87–95.
- JUOLA, J. F., HAUGH, D., TRAST, S., FERRARO, R. and LIENBHABER, M., 1987, Reading with and without eye movements. In J. K. O'Regan and A. Levy-Schoen (eds), *Eye Movements: from Physiology to Cognition* (Amsterdam: Elsevier), 499–508.
- JUOLA, J. F., TIRITOGLU, A. and PLEUNIS, J., 1995, Reading text presented on a small display. *Applied Ergonomics*, **26**, 227–229.
- JUOLA, J. F., WARD, N. J. and McNAMARA, T., 1982, Visual search and reading of rapid serial presentations of letter strings, words, and text. *Journal of Experimental Psychology: General*, **111**, 208–227.
- JUST, M. A. and CARPENTER, P. A., 1980, A theory of reading: from eye fixations to comprehension. *Psychological Review*, **87**, 329–354.
- JUST, M. A., CARPENTER, P. A. and WOOLLEY, J. D., 1982, Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, **111**, 228–238.
- KANG, T. J. and MUTER, P., 1989, Reading dynamically displayed text. *Behaviour and Information Technology*, **8**, 33–42.
- MASSON, M. E. J., 1983, Conceptual processing of text during skimming and rapid sequential reading. *Memory and Cognition*, **11**, 262–274.
- MOORE, D. and ZABRUCKY, K., 1995, Adult age differences in comprehension and memory for computer-displayed and printed text. *Educational Gerontology*, **21**, 139–150.
- MUTER, P., 1996, Interface design and optimization of reading of continuous text. In H. van Oostendorp and S. de Mul (eds), *Cognitive Aspects of Electronic Text Processing* (Norwood, NJ: Ablex), 161–180.
- MUTER, P., KRUK, R. S., BUTTIGIEG, M. A. and KANG, T. J., 1988, Reader-controlled computerized presentation of text. *Human Factors*, **30**, 473–486.
- MUTER, P. and MAURUTTO, P., 1991, Reading and skimming from computer screens and books: the paperless office revisited? *Behaviour and Information Technology*, **10**, 257–266.
- NOORDMAN, L. G. M. and VONK, W., 1997, The different functions of a conjunction in constructing a representation of the discourse. In J. Costermans and M. Fayol (eds), *Processing Interclausal Relationships: Studies in the Production and Comprehension of Text* (Mahwah, NJ: Lawrence Erlbaum Associates), 75–93.
- O'CONNELL, D. C. and KOWAI, S. H., 1986, Use of punctuation for pausing: oral readings by German radio homilists. *Psychological Research*, **48**, 93–98.
- O'HARA, K. and SELLEN, A., 1997, A comparison of reading paper and on-line documents. In *Proceedings of CHI '97*, Atlanta, GA (New York: ACM Press).
- POSTIGO, J. M. L., VALVERDE, G. M. and LLOBELL, J. P., 1992, Analisis y segmentación de textos. *Investigaciones Psicológicas*, **10**, 213–223.
- POTTER, M. C., 1984, Rapid serial visual presentation (RSVP): a method for studying language processing. In D. E. Kieras and M. A. Just (eds), *New Methods in Reading Comprehension Research* (Hillsdale, NJ: Erlbaum), 91–118.
- RAHMAN, T. and MUTER, P., 1999, Designing an interface to optimize reading with small display windows. *Human Factors*, **41**, 106–117.
- REICHER, G. M., 1962, Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, **81**, 275–280.
- ROTHKOPF, E. Z., 1971, Incidental memory for location of information in text. *Journal of Verbal Learning and Verbal Behaviour*, **10**, 608–613.
- RUBIN, G. S. and TURANO, K., 1992, Reading without saccadic eye movements. *Vision Research*, **32**, 895–902.
- STEVENS, J., 1996, *Applied Multivariate Statistics for the Social Sciences*, 3rd edn (Mahwah, NJ: Lawrence Erlbaum Associates).
- STEVENS, S. S. and GALANTER, E. H., 1957, Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, **54**, 377–411.
- STINE, E. A. L., 1990, On-line processing of written text by younger and older adults. *Psychology and Aging*, **5**, 68–78.
- TAYLOR, I., 1976, *Introduction to Psycholinguistics* (New York: Holt, Rinehart & Winston).
- TOWNSEND, D. J., 1983, Thematic processing in sentences and texts. *Cognition*, **13**, 223–261.
- TOWNSEND, D. J., CARRITHERS, C. and BEVER, T. G., 1987, Listening and reading processes in college- and middle school-age readers. In R. Horowitz and S. J. Samuels (eds), *Comprehending Oral and Written Language* (San Diego, CA: Academic Press), 217–242.
- WILLIAMSON, N. L., MUTER, P. and KRUK, R. S., 1986, Computerized presentation of text for the visually handicapped. In E. Hjelmquist and L. G. Nilsson (eds), *Communication and Handicap: Aspects of Psychological Compensation and Technical Aids* (Amsterdam: North Holland), 115–125.