Applying the Rapid Serial Visual Presentation Technique to Small Screens

Karin Sicheritz
karinsi@stp.ling.uu.se

Language Engineering Programme
Department of Linguistics
Uppsala University

Supervisors:
Lars Borin, Uppsala University
Associate Professor Mikael Goldstein,
Ericsson Research, Usability & Interaction Lab, Kista
Abstract

This thesis describes the implementation and the evaluation of a graphical user interface (GUI) employing the Rapid Serial Visual Presentation (RSVP) technique. The GUI was implemented on a Personal Digital Assistant (PDA) featuring a 4.5-inch touch-sensitive colour screen.

The aim of the user evaluation was two-fold: firstly, to compare self-paced RSVP with reading on paper medium, in terms of reading speed, reading comprehension, reading efficiency, experienced workload and readers’ attitude; secondly, to examine whether sonification (adding audio information) of the text is a way to heighten the reading experience when reading with RSVP.

The user evaluation was conducted with ten subjects using a repeated-measurement design. The texts used were six chapters of August Strindberg’s ”Röda Rummet”. Four different conditions were administered: book, two self-paced RSVP formats, and a sonified RSVP format presented at constant speed. The evaluation showed no significant differences in reading speed, comprehension, or reading efficiency between reading from the book and the self-paced RSVP conditions. A linear speed-accuracy trade-off between comprehension and reading speed was found. For each 100-word increase in reading speed comprehension decreased by 9%. The subjects rated all six factors in a workload index, except Physical demand, significantly higher for the self-paced RSVP conditions than for paper medium. Regarding attitude, subjects found RSVP significantly more difficult, less comprehensible, and less stimulating than reading from a book.
Acknowledgment

This work was a joint project between Ericsson Research and Telia Research and was funded by Ericsson Research.

First of all, I would like to thank my academic supervisor at Uppsala University, Lars Borin and my industrial supervisor at Ericsson Research, Usability & Interaction Lab, Associate Professor Mikael Goldstein.

I would also like to thank members of the usability team, Mikael Anneroth and Jost Werdenhoff for technical support, Torbjörn Lundberg at Telia Research for valuable ideas and Niklas Rundquist for composing the soundscape.

Many thanks to fellow students and former fellow students, at the Language Engineering Programme for support and useful ideas: Camilla Bengtsson, Camilla Löfling, Fredrik Olsson, Leif-Jöran Olsson and Per Johnsson.

Last, but not least, many thanks to Liz Osborne at Aisling Ireland Language Services in Dublin for proof-reading this thesis.
# Table of Contents

**Abstract** ................................................. i
**Acknowledgments** .................................... ii
**Table of Contents** ................................. iii
**List of Figures** ..................................... v
**List of Tables** ...................................... v

1 **Introduction** .................................... 1
   1.1 The Assignment .................................. 1
   1.2 Organization of this Thesis ..................... 2

I **Background** .................................. 3

2 **The Reading Process** ......................... 3
   2.1 Identification and Interpretation ............. 3
   2.2 Speed Reading and Skimming .................... 5
   2.3 Reading on Computer Screens .................... 6

3 **What is Readability?** ......................... 7
   3.1 Readability vs. Legibility ....................... 7
   3.2 Readability vs. Comprehensibility .......... 7
   3.3 Readability Measures ............................ 9

4 **Dynamic Text Presentation** ................. 10
   4.1 Scrolling ........................................ 10
   4.2 Leading ........................................ 10
   4.3 Rapid Serial Visual Presentation ............. 11

5 **User Interfaces** ............................... 11
   5.1 Three Paradigms ................................ 12
   5.2 Audio information ............................... 12

6 **Related Work** ................................. 13
   6.1 RSVP ........................................ 13
   6.2 Electronic Books ............................... 15

II **The Prototype** ............................... 16

7 **Graphical Interface** ......................... 16
   7.1 Menus ......................................... 16
   7.2 Control Buttons ................................ 16
   7.3 Text and Page Window ......................... 16

8 **Technical Description** ..................... 17
   8.1 Window Width ................................ 17
   8.2 Speed ......................................... 18
   8.3 Page Window ................................ 18

III **Evaluation** .................................. 19
TABLE OF CONTENTS

9 Method
  9.1 Subjects .................................................. 19
  9.2 Equipment ................................................ 19
    9.2.1 Usability Laboratory .................................. 19
    9.2.2 Texts .................................................. 20
    9.2.3 The PDA and the Application ......................... 20
    9.2.4 Settings .............................................. 20
    9.2.5 Text Sonification .................................... 21
    9.2.6 Comprehension Questionnaire ....................... 21
    9.2.7 Workload index ....................................... 21
    9.2.8 Attitude Inventory .................................. 22
  9.3 Experimental Design .................................... 22
    9.3.1 Null Hypotheses ...................................... 22
    9.3.2 Independent Variables ............................... 23
    9.3.3 Dependent Variables ................................. 23
    9.3.4 Procedure ........................................... 23

10 Results
  10.1 Objective Findings .................................... 25
    10.1.1 Reading Speed ....................................... 25
    10.1.2 Comprehension ....................................... 27
    10.1.3 Reading Efficiency ................................... 27
  10.2 Subjective Findings ................................... 27
    10.2.1 Workload ............................................. 27
    10.2.2 Attitude ............................................. 29

11 Discussion .................................................. 30

IV Concluding Remarks ....................................... 34

12 Further Work ................................................ 34

13 Summary ...................................................... 34

References ........................................................ 35

A Questionnaire Chapter 1 .................................... 39
B Questionnaire Chapter 3 .................................... 40
C Questionnaire Chapter 4 .................................... 41
D Questionnaire Chapter 6 .................................... 42
E Workload inventory .......................................... 43
F Attitude inventory ........................................... 44
LIST OF FIGURES

List of Figures

1. Rocket eBook. ........................................ 15
2. The GUI in full size. ................................... 17
3. The prototype Reader 1.0 implemented on a Casio Cassiopeia E–105. ................. 21
4. Box plot presentation of median, 25%–75% percentile and min-max values for reading speed for the self-paced conditions Book, RSVP11 and RSVP25. ............... 26
5. Box plot presentation of median, 25%-75% percentile and min-max values for reading efficiency (wpm) for all four conditions. ............................. 28
6. Plotting comprehension (%) as a function of reading speed (wpm) for the three self-paced conditions showed a significant negative correlation (R=-0.452, p=0.012). .......... 29
7. Ratings (in millimetre) of workload for condition Book, RSVP11, and RSVP25 using the NASA-TLX workload index. Low ratings indicate low workload and good performance. ............................... 30

List of Tables

1. Reading goal and connected reading process. ........................................ 8
2. Reading goal and connected text properties. ........................................ 8
3. Interpretation of lix. ........................................... 9
4. Text properties with regard to number of words and lix for the six initial chapters of “Röda Rumnet”. ......................................................... 20
5. Overview of the experimental design with respect to text properties, conditions, dependent variables and procedure according to chapter. .......................... 25
6. Minimum, maximum, mean and median values for reading speed (wpm) for the self-paced conditions. ......................................................... 26
7. Average and median values for reading efficiency (wpm) for the self-paced conditions. 27
1 Introduction

The expansion of computer technology in the last decades has given birth to a new format for written communication, namely, electronic documents. Texts, such as online journals and electronic mail, are used by many people in everyday life. Reading on a computer screen differs in many ways from reading on paper and several studies have shown that reading speed is 15%–30% slower for reading computer-presented text than for reading text presented on paper (Edeborg 1999; Mills & Weldon 1987; Shneiderman 1998).

Since the paper medium has been the dominant one for centuries this way of reading is what people and society are adapted to. Many readers find it tiring to read on a computer screen and often print electronic documents before reading them. This is a waste of natural resources as well as time since the text of a letter, a journal or a novel can be regarded more as a service than as a device; the information is consumed at the same moment as it is presented.

Another problem arises with hand-held devices, such as PDAs and mobile phones, with a limited screen space. Reading on small displays is an important issue given the large volumes of hand-held devices now in use. In the present situation it is possible to download, and store, a number of novels on a PDA. Devices designed specially for the purpose of reading are now available on the market. Examples of such devices, often called electronic books, are the Rocket eBook (Rocket eBook 2000) and SoftBook (SoftBook 2000). Typically, the text is displayed in a conventional way, i.e. on pages. This implies the use of scroll buttons, for vertical scrolling, or turn-page keys to move forward in the text.

In the conventional way of reading, both on paper and on screens, the eye has to traverse the text which is presented spatially. Cognitive load is used for eye movements and planning for next gaze, which may impair reading speed and comprehension. This traditional way of presenting the text is called static text presentation. The counterpart, dynamic text presentation, is feasible in the electronic medium. Utilizing the special capabilities of the computer to display text in ways other than static presentation has become more urgent as screens in everyday use have decreased in size.

1.1 The Assignment

The purpose of this assignment was to implement and to evaluate a graphical user interface (GUI) employing the Rapid Serial Visual Presentation (RSVP) technique for text presentation on a Personal Digital Assistant (PDA). The evaluation was conducted in a user study where reading with the GUI was compared to reading on paper. Other parts of the assignment were to design the study, to carry it out and to analyze the results.
1.2 Organization of this Thesis

This thesis is organized into four main parts. The first part gives the theoretical background related to this work. A description of the GUI Reader 1.0 is the topic of the second part. The third part describes the user study of the Reader 1.0 and part four concludes the thesis with proposals for further work in this area and a summary.
Part I

Background

This part of the thesis starts off in section 2 with background information concerning the reading process from a general point of view. Section 3 gives a brief description of the concept of readability. Up to this point in the thesis it is assumed that text presentation is static. The counterpart to static text presentation, i.e. dynamic text presentation, is described in section 4. Some topics concerning user interfaces are dealt with in section 5. Section 6 concludes this part with a description of earlier work in this area.

2 The Reading Process

The reading process is defined in Robeck and Wallace (1990:27) as follows.

Reading is a process of translating signs and symbols into meanings and incorporating the new information into existing cognitive and affective structures.

The study of reading involves different disciplines like pedagogy, psychology, linguistics and neurology. There is a huge variation in reading modes and the reader’s goal is one of the factors determining the reading process. Other important determiners are the text and the reading situation (Just & Carpenter 1980; Gunnarsson 1982).

The following sections deal with identification and interpretation of text items (section 2.1) and different types of speed reading (section 2.2).

2.1 Identification and Interpretation

The identification of smaller text units like letters and words is both a visual and a cognitive process which is in itself complex. Massaro et al. (1980) proposed a stage model which provides a framework from which they derive hypotheses about word recognition. In summary, their model consists of a number of stages between the printed word and its meaning. Each stage has a memory component and a process component and the memory components includes both short-term\(^1\) and long-term memory (e.g. Dix 1998). The processing of words and larger text units involves knowledge of orthography, phonology, syntax, and semantics and is regarded as flexible in terms of processing order by a majority of researchers (Gunnarsson 1982).

Reading statically-presented text consists of three distinct visual processes. The perception of the text occurs during pauses, called fixations or gazes. The

\(^1\)Also called working memory.
short, rapid eye jumps between fixations are called \textit{saccades} and the eye movements to the next line are called \textit{return sweeps} (Robeck & Wallace 1990). Before the publication in 1878 of the French ophthalmologist Emile Javal’s articles, the belief was that the eye moved continuously along a line and fixated every letter. Javal’s findings showed that the eyes process several letters and words per fixation and later studies, carried out by J. M. Catell and E. B. Huey among others, showed that, if the word is already known to the reader, reading is rather a processing of a word image than a sequence of letters (Mullen 1997). Catell (1885 referred to e.g. in Gunnarsson 1982) also found that readers were able to identify four to five words forming a meaningful sequence, during a certain time of exposure, but only two words in an incoherent sequence. These results were confirmed in a study, reported in Masson and Borowsky (1998), showing how context influences word identification. The perceptual span has been found to be centered to the right of the point of regard, at least for readers of left-to-right languages (Just & Carpenter 1980) and ranges on average four characters to the left and 12 characters to the right (Robeck & Wallace 1990).

The duration of the saccade, which can be both backwards and forwards, is too short for sufficient processing of the text to occur (Massaro et al. 1980). The backward saccades, called \textit{regressions} constitute 19\% of the saccades and both types of saccades, (backward and forward) constitute about 5\%–10\% of the total reading time (Just & Carpenter 1980).

Although the time needed for the visual processing in a word recognition task was 50 milliseconds (msec) (Carr 1981 referred to in Young 1984) the duration of the fixation is actually more than four times longer. The fixation time is roughly ten times longer than the duration of the saccade, and averages 239 msec (Just & Carpenter 1980). Just and Carpenter tested students reading scientific texts, and reported a large variation in the duration of individual fixations. Their study showed that almost every content word was fixated at least once and that fixation time was longer on words that were infrequent, thematically important, and those words that clarify the interpretation of previous words. Thus most of the gaze duration time is used for integration, inference making, and connecting incoming information with world knowledge (Carr 1981 referred to in Young 1984).

Just and Carpenter also found the gaze duration to be longer at the end of sentences, indicating a greater processing load at these points. The reader searches for referents that have not yet been assigned, integrates information and makes inferences to capture the coherence of the text. One explanation for the longer fixation time is that sentence ambiguity is solved when the sentence is completed. However, Bever et al. (1973) reported no increase in comprehension time when testing complete, ambiguous clauses. Other explanations for increased processing time are that sentence boundaries unambiguously signal the end of one thought.

So-called inner speech occurs in several mental processes and reading is no exception. Huey (1908 referred to e.g. in Robeck & Wallace 1990) found that most readers verbalize the printed symbols into inner speech. It was later found that skilled readers have a lower activity in the muscles activated in speech, than
less skilled readers. Moreover, the activity increases when more difficult texts are read (Platzack 1974).

Familiarity with the topic is naturally an advantage for comprehension and interpretation. Despite this fact few researchers take into account prior knowledge and topic familiarity when investigating other reading comprehension variables (Spyridakis & Wenger 1991).

### 2.2 Speed Reading and Skimming

As mentioned above there are several modes of reading and two of them are *skimming* and *speed reading*. Average “normal” reading rates are about 200–300 words per minute (wpm) (Masson 1985).

The term *speed reading* is used for ‘very fast reading of an entire text with optimal assimilation’. The eye movements seem to distinguish speed-readers from typical readers. Fast readers make fewer fixations and fixation durations are shorter than for poorer readers (Mullen 1997). Masson (1985) estimated the maximal reading rate at about 800–900 wpm and reported that reading beyond these rates would imply skipping many words. Reading efficiency, i.e. error-corrected reading speed, has been used as an attempt to avoid problems associated with speed-accuracy trade-off\(^2\) (Wickens 1992). Reading efficiency is measured as reading speed times percentage correct on a comprehension test (Jackson & McClelland 1979; Rahman & Muter 1999). There are wide differences among readers with regard to speed accuracy trade-off and free reading rate is not consistently correlated with comprehension performance (Masson 1985). Masson (1985:189-196) proposed four components that might contribute to differences in the speed-accuracy trade-off:

- Visual span of perception.
- Speed of access to memory codes.
- Working memory.
- General language comprehension ability.

*Skimming* is most often used meaning ‘rapid moves through text to locate particular information or to get a general sense of the material’. A great amount of cognitive capacity is used in this type of reading (Robeck & Wallace 1990; Muter 1996). In a study reported by Masson (1983) where speed-readers were compared to skimmers and normal readers, speed-readers had an average reading rate of 700 wpm, skimmers read at 600 wpm and normal readers on average 240 wpm. Masson found that normal readers had an advantage in comprehension tests and the only advantage for speed-readers was for easier narrative material.

\(^2\)As speed increases performance decreases.
2.3 Reading on Computer Screens

As mentioned in previous section normal reading on paper averages 200–300 wpm. A number of studies have shown that proof-reading on a Cathode Ray Tube (CRT) is 20%–30% slower than reading on paper (Gould & Grischkowsky 1984; Muter et al. 1982; Edeborg 1999). On the other hand, Gould and Grischkowsky found no evidence for decreased reading comprehension or proof-reading accuracy when proof-reading on screen compared to paper.

Shneiderman (1998:412–413) outlines some disadvantages in computer reading compared to paper reviewing ten studies which showed 15–30% slower reading speed. These studies employed large screens.

- Fonts may be poor, especially on low resolution displays
- Low contrast between characters and background
- Emitted light may be more difficult to read, flicker may be a problem as well as the curved display surface
- Small displays require frequent page turning
- Reading distance can be greater and display placement may be too high for comfortable reading
- Layout and formatting problems such as improper margins, inappropriate line width (35–55 characters are recommended), or awkward justification (left justification and ragged right is recommended)
- Reduced hand and body motion as well as rigid body posture may be tiring
- Unfamiliarity with displays

Several factors affecting reading on screen were outlined in Edeborg (1999). Some of the variables having an adverse effect on reading speed were found to be resolution, edge sharpness, contrast, distance between the reader and the material and body posture. Edeborg also found that the font size had an impact in on-screen reading but that there is no difference in reading speed between fonts with and without serif.

In contrast, Muter and Maurutto (1991) found no significant difference in reading efficiency for normal reading rates from book and screen. However, skimmers performed better reading from a book than from a screen in terms of reading efficiency due to slower reading rates on screen. The difference may have been a result of a speed-accuracy trade-off since comprehension was higher on screen compared to book.

---

3The predominant display device, similar to a television monitor. An electronic beam sweeps out lines of dots to form letters and graphics. CRT sizes (measured diagonally) range from less than 2 inches to more than 30 inches; popular models range 11 to 17 inches (Dix et al. 1998; Shneiderman 1998).
Background color, texture, lightness were tested in Scharff et al. (1999) using black text on three background colors, and within each color, two lightness settings. Results obtained showed an advantage to plain, yellow and grey backgrounds, over blue, textured backgrounds. Lightness did not seem to affect reading speed in this experiment.

3 What is Readability?

*Readability* has different interpretations in literature and this section is an attempt to clarify matters respecting this concept.

*Readability* is most often used to mean ‘easy to read’ in relation to the text style or, put in another way, the effort made by the reader to understand the content of a text (Gunnarsson 1982). Different texts are, of course, judged differently by different readers and as Gunnarsson (1982:51) pointed out: “Readers differ, not only in their world knowledge, but also in their ability to interpret the text from a linguistic point of view”.

Readability deals with the linguistics qualities of a text and is measured in tasks which require reading continuous meaningful texts. Readability is typically measured by reading comprehension and reading speed (Mills & Weldon 1987). Other objective productivity measures are reading efficiency (Jackson & McClelland 1979; Rahman & Muter 1999) and proof-reading accuracy (Gould & Grischkowsky 1984). Several factors are involved in deciding whether a text is easy or hard to read.

The following sections compare readability to *legibility* and *comprehensibility* and a brief description is given of *readability measures*.

3.1 Readability vs. Legibility

Readability and legibility are defined by Mills and Weldon (1987:331) as follows:

Readability generally refers to the ease with which the meaning of the text can be comprehended, whereas legibility generally refers to the ease of identification of the text items.

Legibility deals with typographical factors, for instance characters, formatting, contrast and color, and is generally measured by means of identification tasks in which single letters or a smaller amount of letters are presented in visual noise (Mills & Weldon 1987). Thus, legibility is restricted to typographical factors.

3.2 Readability vs. Comprehensibility

Gunnarsson (1982) makes a distinction between readability and comprehensibility. This distinction is based upon the most common ways to measure readability
and on existing readability formulas (see section 3.3). Gunnarsson claims that the readability, measured in comprehension tests, refers to comprehension on a shallow level and no attention is paid to the reader and the specific reading situation. On the contrary, a high score on these tests may be a sign of not understanding, since most of the tests measure recall of details. Comprehensibility, on the other hand, take into account the reader and the reading goals. Gunnarsson proposes a model which connects different text properties and reading processes to different reading goals. The reading goals can be interpreted as different levels of comprehension. The reading goals (or levels of comprehension) and their connection to certain reading processes and text properties are outlined in table 1 (Gunnarsson 1982:70) and table 2 (Gunnarsson 1982:82) respectively.

<table>
<thead>
<tr>
<th>Reading goal</th>
<th>Reading process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization of the text surface</td>
<td>Identification of words</td>
</tr>
<tr>
<td>Registration of the content</td>
<td>Relating the text to the message</td>
</tr>
<tr>
<td>Comprehension of the writers perspective</td>
<td>Relating the text to the writer’s intention</td>
</tr>
<tr>
<td>Integration in own reality</td>
<td>Relating the text to the reader’s own world knowledge</td>
</tr>
<tr>
<td>Action-directed comprehension</td>
<td>Relating the text to the reader’s own situation and to different action alternatives</td>
</tr>
</tbody>
</table>

Table 1: Reading goal and connected reading process.

<table>
<thead>
<tr>
<th>Reading goal</th>
<th>Text properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization of the text surface</td>
<td>Typography, word semantics, syntax</td>
</tr>
<tr>
<td>Registration of the content</td>
<td>Syntax, clause semantics</td>
</tr>
<tr>
<td>Comprehension of the writers perspective</td>
<td>Perspective</td>
</tr>
<tr>
<td>Integration in own reality</td>
<td>Perspective</td>
</tr>
<tr>
<td>Action-directed comprehension</td>
<td>Perspective, function</td>
</tr>
</tbody>
</table>

Table 2: Reading goal and connected text properties.

Almlöv (1991) examined the comprehensibility of a schoolbook by a number of text linguistic analyses. The schoolbook in focus in this study was criticized by many students who judged it to be difficult to read. The results of the analyses showed a lack of good thematic structure, weak reference bindings, and lack of connections between clauses which might have been the reasons for the students’ dissatisfaction with the book.

The major reason for the difficulty a of text is, according to Platzack (1987), a combination of difficult words and the words’ level of abstraction.

In the fifties Benjamin Bloom defined a taxonomy regarding learning in a quantitative perspective. The taxonomy constitutes six levels, arranged in hierarchical order. Thus, to cope with a new level, one has to master the lower levels. The first level, called Knowledge differs from the higher levels in that the psychological process involved is remembering. Higher levels involve processes which are more complex, for instance, relating and reorganizing. The following are brief de-
3 WHAT IS READABILITY?

Descriptions of the levels in Bloom’s Cognitive Taxonomy (Bloom 1956).

1.00 Knowledge: To recognize accurate statements in response to particular questions, to learn facts, a literal direct recall of information.

2.00 Comprehension: To translate, interpret and extrapolate the objective of the material.

3.00 Application: To apply the objectives of the material given in a more generalized manner, into a concrete level of abstraction, to solve problems.

4.00 Analysis: To dissemble the material and find relationships between its parts.

5.00 Synthesis: To produce unique communication, to produce a plan, and to derive a set of abstract relations.

6.00 Evaluation: To make a judgment of the given material.

3.3 Readability Measures

As mentioned above readability refers to understanding on a shallow level. Nevertheless there has been a need to predict the readability of texts in different areas, for instance in the business world, in the industry and in teaching. Several readability measures have been constructed and linguistic factors often used in these formulas are the number of different word types, the number of infrequent words, word length in syllables (McLauaglin 1969) or letters (Björnsson 1968), the number of words per sentence etc. Thus, the formulas are based on quantitative and mechanic measures. Most of the formulas include a word factor and a sentence factor (Gunnarsson 1982) while the factors mentioned in section 3.2 are not reflected, for example the syntax or the level of abstraction.

Lix (in Swedish läsbarhetsindex) is a readability measure for Swedish texts developed by Björnsson (1968). The formula is based upon the percentage of long words (six letters or more) and the average sentence length (words per sentence). The lix formula is as follows.

\[ \text{lix} = \text{sentence length} + \text{long words} \]

The interpretation of lix values is outlined in table 3 (Björnsson 1968).

<table>
<thead>
<tr>
<th>Lix</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>Very easy</td>
</tr>
<tr>
<td>30</td>
<td>Easy</td>
</tr>
<tr>
<td>40</td>
<td>Medium</td>
</tr>
<tr>
<td>50</td>
<td>Difficult</td>
</tr>
<tr>
<td>60</td>
<td>Very difficult</td>
</tr>
</tbody>
</table>

Table 3: Interpretation of lix.

The different text types tested by Björnsson were children’s books, fiction, technical literature and newspaper articles. The average lix value for 100 fiction books was found to be 33 and 95% of the material ranged from 28 to 38.
Findings that are contradictory to the idea that longer sentences are less readable are reported by Platzack (1974) where shorter sentences (on average nine words) were less readable than longer sentences (13 words), measured as the readers’ recall of details.

4 Dynamic Text Presentation

Dynamic text presentation has been used in several studies since the seventies as an attempt to optimize reading on screens. It has been shown that reading dynamically displayed text is as efficient, in terms of reading speed and comprehension, as reading text presented statically on a screen (Rahman and Muter 1999). In the following sections three techniques are described: vertical scrolling in section 4.1, leading in section 4.2 and in section 4.3 the Rapid Serial Visual Presentation technique.

4.1 Scrolling

Text layout on the computer screen is usually presented according to the paper metaphor (see section 5.1). The only difference is that instead of paper page units, a scroll bar or a turn-page key is used to move forward in the text. Studies of vertical scrolling reported by Mills and Weldon (1987) indicate that continuous scrolling is preferred to discrete, line-by-line, scrolling and that scrolling an entire block of text is better than scrolling individual lines or sentences. It is also reported that paging, that is, paged information changed all at once, was preferred to continuous scrolling. The latter result was obtained with inexperienced users which may have influenced that result.

4.2 Leading

One dynamic text presentation is leading, also known as the Times Square technique named after the location where the most famous implementation of this technique was found. In leading, the text is scrolled horizontally from right to left and is often seen on one line displays.

In a study by Kang and Muter (1989) three leading formats were tested; one pixel at a time, one letter at a time and one word at a time. These formats were compared with the RSVP technique (see section 6.1) at presentation rates ranging from 100 to 300 wpm. The results showed that the leading conditions were as good as the RSVP technique, inconsistent with results reported by Grannas et al. (1984) who argued that leading was not a viable format when presenting text at rates approaching normal reading speed. Joula et al. (1995) found on the contrary leading to be inferior to RSVP when testing texts from a semi-scientific magazine. The reading conditions were tested at two rates (171 and 260 wpm) and in lower-case fonts (the initial letter of the first word in each sentence was capitalized) or all
letters in upper-case. The sentences, averaging 10 words in length, were presented on an eight-character display.

In Chen and Chan (1990) user-controlled leading rate was compared to experimenter-controlled rate. Forty-eight subjects were tested in four 20-minute sessions run in four consecutive days. The results showed, for the self-paced subjects, increased performance with regard to reading speed over the time of the study but no improvements with regard to comprehension. For subjects reading with experimenter-controlled rates comprehension increased over time showing that experience is an important factor in reading moving text.

4.3 Rapid Serial Visual Presentation

The Rapid Serial Visual Presentation technique (RSVP) simulates the visual experience of normal reading but eliminates the need for eye movements. The term RSVP was first introduced by Forster (1970) who experimented with single words presented at a fixed location on the screen but the concept had been used more than ten years earlier by Gilbert (1959 referred to in e.g. Granaas et al. 1984). A variant on RSVP, called moving RSVP, presents small text segments (one or a few words) successively but the location is in the position it would occupy if the entire text was displayed on the screen. In this thesis the term RSVP is referred to as stationary RSVP and the definition of RSVP used in this thesis would be: successive text presentation in small segments, in a fixed location, on a single line for a preselected amount of time.

One consequence of reading with the RSVP technique is that the number of eye movements is minimized, especially if the text units are as small as single words. In addition, regressions are effectively prevented. With this technique the need to devote attention to planning the next gaze is eliminated since the gaze is directed at a fixed location. Another consequence is that the reader cannot easily skip down lines and therefore may be able to recall more information than in normal reading. On the other hand, if the text units consist of a single word, the reader may have to make a greater effort to determine which words form a coherent meaning unit (Just & Carpenter 1980).

RSVP has been used as a tool for studying reading behavior (Forster 1970; Juola et al. 1982). Proposed application areas for this text presentation technique are, for instance, use by visually impaired readers (Fine & Peli 1995) and when display space is limited (Granaas et al. 1984; Muter 1996). An implementation of the RSVP technique for large screens is available on the Internet (Vortex 1998).

5 User Interfaces

A user interface has to be judged within a context which, naturally, is focused on the user. Examples of human factors measured in the evaluation of user interfaces are the time to learn tasks, the speed of performance, how many and what kind of
errors the users make, and subjective satisfaction (Shneiderman 1998).

Two topics concerning the user interface described in this thesis (see part II) are outlined in the following two sections.

5.1 Three Paradigms

Alan Cooper (1995) makes a distinction between three dominant paradigms respecting interface design from the user’s point of view. These are the technology paradigm, the metaphor paradigm and the idiomatic paradigm.

The technology paradigm is based on the understanding of how the software works, i.e. the user needs to be familiar with the structure of the program.

The metaphor paradigm, upon which many interfaces are based, relies on the user’s intuition and is a step forward from the technology paradigm. However, Cooper claims that the ability to recognize the metaphors is questionable due to, for instance, cultural differences. An example of this paradigm is the desk-top metaphor whose functionality is immediately recognized by the user.

The idiomatic paradigm has to be learned and as Cooper (1995:59) pointed out: “All idioms must be learned. Good idioms only need to be learned once.” The distinction between the technology and the idiomatic paradigm is that the latter implies learning without the need to understand.

5.2 Audio information

An alternative mode of output, used in conjunction with the computer screen, is auditory signals. It is not yet known how to use sounds to achieve maximum effect in interactive systems and besides in video games, the auditory channel is comparatively little used in standard interfaces (Dix et al. 1998).

Auditory icons (Gaver 1986) use natural sounds to represent different types of objects and actions in the interface. These are everyday sounds meant to convey information about events in the computer by analogy with everyday events. Gaver classified the mappings between data and their auditory representation into three different types: symbolic, nomic and metaphorical. Symbolic mappings rely on social convention for meaning, such as applause for approval. Nomic representations are physical, such as the sound of a closing metal cabinet for closing a file. Metaphorical mappings represent similarities, such as a falling pitch for a falling object.

The term earcons is sometimes used for synthetic sounds (Dix et al. 1998) but also in a more general sense as the audio counterpart to icons (Blattner et al. 1989).

A method for linking nomic auditory icons and earcons to a passage of text presented in the RSVP paradigm is proposed (Goldstein et al. 2000). In order to heighten the reading experience the soundscape is played back to the reader when the linked text is displayed at any selected reading speed.
6 Related Work

Section 6.1 below deals with some experiments conducted on the RSVP paradigm and section 6.2 outlines a couple of devices specially designed for reading electronic fictional texts.

6.1 RSVP

The RSVP technique has been explored in a number of studies since the seventies. The output device in these experiments has been a CRT terminal and the texts used were semi-scientific or artificial texts ranging from 1400 to 8700 words. Typical subjects in studies reviewed here were young university students with one exception, namely in Fine and Peli (1995), where elderly people participated.

Joula et al. (1982) compared reading of shorter paragraphs in two conditions; some presented in their entirety on a CRT screen and some presented via the RSVP technique. In the RSVP condition the text segments consisted of 5, 10 and 15 characters and were presented with a duration of 200 and 300 msec each which means reading rates ranging from about 200 to 700 wpm. Forty-eight subjects, all university students, were tested and each paragraph was followed by multiple-choice questions testing memory for both specific details and more general information as well as logic inferences. The experiment showed no significant differences in comprehension between the two reading conditions.

In a study performed by Cocklin et al. (1984) a variety of segment sizes and durations were used to investigate the optimal segment length (in characters) for RSVP. The text difficulties ranged from “intermediate” to “secondary” and 72 subjects were tested, all of them university students. Comprehension was tested by multiple-choice questions. The average window size varied between five and 20 characters and reading rates were 200, 400, 600 and 800 wpm. Results showed that comprehension was found to peak at segment lengths averaging 12 characters across the range of reading speeds and text difficulties. The findings in this study indicate a limit of perceptual capacity in combination with the need for context while reading.

Studies using clause-sized text segments presented at a rapid rate have shown that RSVP is superior to reading statically presented text on a screen measured as recall on comprehension test (Cocklin et al 1984; Young 1984). The reasons for these findings were proposed to be that clauses approximate the chunks the reader constructs, and although eye movements are not eliminated, the number of regressions are reduced, as well as the tendency to skip down lines.

Rahman and Muter (1999) reported no significant differences concerning comprehension and reading speed between statically presented text on screen and two RSVP formats. Output device in these experiments was a 15-inch CRT screen. The two RSVP formats were used either alone or with a completion meter. Participants were instructed to read as quickly and as accurately as possible. The RSVP formats required the participants to press a key to advance to the next sentence. One
passage of text, averaging 400 words, was used in each of the five conditions. The tests were preceded by practice sessions reading about 1200 words in each condition. All subjects performed in all conditions. The results showed no significant differences regarding comprehension and reading speed.

Fine and Peli (1995) compared visually impaired to normally sighted, elderly subjects (average age was 72.1 years) reading in two conditions: scroll and RSVP. They reported similar reading rates for scrolled and RSVP texts for the visually impaired. Only subjects with normal vision showed benefit from the RSVP condition.

Comprehension has been shown to increase when sentences are presented in linguistically appropriate text segments and two studies that confirm this idea were carried out in the eighties. Text segmentation based on linguistic structure was explored by Pynte and Noizet (1980) to determine the optimal segmentation for French sentences. This segmentation was compared to word-by-word presentation and no segmentation. The results showed that the fastest reading time was obtained for sentences presented in segments consisting of phrase constituents. For short sentences (seven or eight words) the noun phrase was separated from the verb phrase. For longer sentences (nine to 12 words) the segment could also consist of e.g. adjective-noun pairs or prepositional phrases. The reading time was found to increase greatly when syntactically linked words were separated.

Cocklin et al (1984) tested 40 students in an experiment using paragraphs in English segmented in “idea units” consisting of two or three words. These “idea units” were compared to segments averaging 13 characters, rounding the number of characters to the nearest word. The “idea units” were not all linguistically defined components such as segmented on clause and phrase boundaries. The segmentation was done by four independent judges who reached about 90% agreement. The reading speed averaged 300 wpm. Results showed a perceptual advantage to reading text divided in “idea units” though comprehension increase was not significant.

A pause, 500 or 1000 msec in duration, inserted between sentences were found to significantly improve RSVP performance in a study with RSVP and skimming tasks (Masson 1983). The duration of the blank window had no effect. Muter (1996) recommended a blank window 250–500 msec in duration.

In summary the studies above showed the optimal conditions for the RSVP paradigm would be as outlined below:

- On average 12 characters in window width
- Two or three words per segment (English)
- Linguistically coherent segments
- A blank window should be inserted between sentences
6.2 Electronic Books

Two electronic books available on the market are for instance the *Rocket eBook* (Rocket eBook 2000), and the *SoftBook* (SoftBook 2000). Both these devices are designed according to the book metaphor and the text is presented in the traditional way.

*Rocket eBook* (see figure 1) has a touch sensitive 11.5 x 8 cm screen with a resolution of 480 x 330 pixels, and a back light to allow reading in darkness. This device permits setting bookmarks, underlining the text displayed, adding notes and in addition, a dictionary is available. *Rocket eBook* needs to be connected to a stationary computer for downloading of texts. The location of the current page is displayed as a percentage of the total length. *Rocket eBook* can store ten average novels (3 200 pages), has a battery life of 20–40 hours and weighs about 700 grams.

*SoftBook* features a 15 x 20 cm (9.5–inch), grey scale, back lighting, touch sensitive screen. This device weighs 1.3 kg, has a built-in modem and features tools for bookmarking, hyper linking and text mark-up. *SoftBook* has a storage capacity of approximately 4 000 pages.
Part II
The Prototype

In order to evaluate the RSVP technique a prototype, Reader 1.0 was implemented by the author. Using the Personal Java environment the prototype was implemented on a Casio Cassiopeia E–105 (described in section 9.2) which is based on WindowsCE. The program takes a text file in ASCII format as input. The text is divided into segments which are temporally displayed in a text area at a certain rate. The user can choose the window width (the length of the text segment in characters) and the speed level. The following sections present the GUI and give a technical description of the application.

7 Graphical Interface

The GUI, shown in figure 2, consists of a menu bar and a label presenting the title of the text currently being read. Furthermore, there are soft buttons for starting, continuing, pausing and looking back at the previous presented text portion. In the centre of the GUI there is a text window where the running text is displayed, and at the bottom of the GUI there is a textual completion meter. The latter displays the page currently read on the left and the total number of pages for the current text on the right.

7.1 Menus

The menu bar at the top of the GUI consists of one menu for each of the following categories: the text to be read, window width, speed level and font size. There are currently three different window widths (11, 17 and 25 characters), twenty-seven speed levels ranging from approximately 25 to 1000 wpm (low, 1–25 to fast), and five font sizes (8–16 points). The chosen settings for window width, speed level and font size are displayed beneath the menu bar. The user can change these settings at any time while reading.

7.2 Control Buttons

The GUI has four buttons for interactive control: Start for initiating the text run, Pause for pausing, Continue for continuing to read (after a pause) and Previous for viewing the previous and the current text unit at the same time. When a text has been chosen the title is displayed to the right of the start button.

7.3 Text and Page Window

The text window consists of an area for presentation of each text segment. The text segment is presented left adjusted and vertically centered. The page window
presents the total number of pages of the current text and the current page being read.

8 Technical Description

The program consists of two files, one for the graphical interface and for control of the user actions performed, and one for the functions that run the program. The following three sections describe how the text is divided in segments, the calculation of the presentation rate and how the page numbering is calculated.

8.1 Window Width

The segmentation of the text works as follows. The program searches for the nearest whitespace given the window width chosen by the user and breaks the character string at this point. If the number of characters for a single word exceeds the chosen window width, the program does not hyphenate but presents the word in full. The program also breaks the character string when commas or other punctuation marks are detected.
A Swedish word has on average 5.34 letters according to Dahlqvist (1999). This figure was calculated by examining a corpus containing articles published in 1995 and 1996 by the Swedish newspapers *Svenska Dagbladet* and *Upsala Nya Tidning*. This corpus consists of more than 70 million tokens and 1.5 million types. Since words are delimited by at least one character (whitespace or punctuation mark plus whitespace), one character was added to the average word length, that is 6.34 characters for each word. Given that every word needs the space of six characters except the last word in the segment since this word does not have to be separated from the following word. Thus, a window width of 11 characters allows for two words per segment and one of 17 characters allow for three words. The 25-character window width is the longest string of characters allowed by the screen space and this width contains four average Swedish words.

A blank window 250 msec in duration appears between sentences in order to improve comprehension.

### 8.2 Speed

As mentioned above punctuation marks, except commas, are followed by a blank window. The fact that the window width is approximate and that punctuation marks are not counted in advance means that the speed will be approximate. The actual speed depends on the chosen window width and the text. The speed levels are discrete steps of approximately 30–40 wpm between each level.

### 8.3 Page Window

Rahman and Muter (1999) used a graphical completion meter in their study and an alternative to this is the textual feedback used here. The current page and the total number of pages are displayed according to the book metaphor. The total number of pages is calculated in such way that every word is counted and the total number of words is divided by 300, i.e. every page contains 300 words.
Part III
Evaluation

In order to benchmark the RSVP technique on a small display against conventional book reading, four different conditions were used: self-paced reading from book (Book), self-paced RSVP with a window width of 11 characters (RSVP11), self-paced RSVP with a window width of 25 characters (RSVP25), and sonified RSVP25 at constant reading pace (156 wpm). Each subject participated in all conditions and read the chapters of the book in consecutive order. Dependent measures were reading speed, comprehension, reading efficiency, workload index (NASA-TLX) and an attitude inventory. The experiment was carried out at a usability laboratory at the department of Design & Interaction Lab at Ericsson Research in Kista between the 2 and 12 May 2000.

In the following sections the method is described (section 9), as well as the results (section 10), and a discussion (section 11) of the obtained results.

9 Method

This section outlines the method used in this study. In the first section the participants in the study are presented, the second section describes the material and equipment used and the third section presents the experimental design.

9.1 Subjects

Ten subjects, five male and five female, participated in the study. The age ranged between 26 and 43 years with an average of 32 years. Selection criteria were Swedish as native language, some experience in using PDAs and an interest in reading fiction. The subjects’ experience of PDAs varied from a little experience to everyday use. The subjects’ reading habits with regard to fiction varied from approximately three to 12 novels per year. None of the subjects had read Strindberg’s “Röda Rummet” before the experiment nor had they used the RSVP technique. All of the subjects received four cinema tickets and a free lunch for participating in the study.

9.2 Equipment

9.2.1 Usability Laboratory

The experiment took place in a usability laboratory which has two rooms, one control room and one test room. These rooms, which are separated by a one-way mirror and are soundproofed, allow the experimenter to observe the subject during the experiment without disturbing him/her. The test room is equipped with microphones and video cameras and the control room has video-capture equipment,
computers and a microphone, the latter for communication with the subject during the experiment.

9.2.2 Texts

The texts used in the experiment were the first six chapters from the novel “Röda Rummet” by August Strindberg in Swedish. These chapters contain 25,360 words in total. Subjects read chapter one from a book and chapter two, three, four, five and six on the PDA. The novel is available in electronic format through Project Runeberg (Project Runeberg 1992).

In table 4 below the computed lix for each chapter and the number of words are listed:

<table>
<thead>
<tr>
<th>Chapter 1</th>
<th>Chapter 2</th>
<th>Chapter 3</th>
<th>Chapter 4</th>
<th>Chapter 5</th>
<th>Chapter 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of words</td>
<td>3971</td>
<td>2724</td>
<td>4120</td>
<td>6309</td>
<td>3182</td>
</tr>
<tr>
<td>Lix</td>
<td>54</td>
<td>51</td>
<td>49</td>
<td>42</td>
<td>38</td>
</tr>
</tbody>
</table>

Table 4: Text properties with regard to number of words and lix for the six initial chapters of “Röda Rummet”.

9.2.3 The PDA and the Application

The application (described in part II) was implemented on a Casio Cassiopeia E–105 (see figure 3). This PDA features a touch-sensitive 4.5–inch color liquid-crystal display\(^5\) with a resolution of 320 x 240 pixels\(^6\) in portrait mode. The battery life is approximately five to six hours and the device weighs 225 grams. A stylus is used as a pointer and the user interacts by tapping on the screen. Since the control buttons on the front side below the display were not used by the GUI, those buttons were covered by a piece of cardboard to prevent subjects from accidently pushing them.

9.2.4 Settings

Chapter one was read from a book (Strindberg 1879). Font was Times, size approximately 12 points. Each page contained approximately 300 words.

The font used in RSVP conditions across the experiment was Courier, 12 points. The text was displayed in black on a white background.

\(^5\)Liquid-crystal displays (LCD) are found on portable computers, watches and calculators because of their thin form, light weight, and low power consumption. Voltage changes influence the polarization of tiny capsules of liquid crystals. LCDs are flicker-free but the resolution is limited due to the size of the capsules (Dix et al. 1998; Shneiderman 1998).

\(^6\)The diagonal of a large computer screen measures 12–20 inches with a resolution of at least 1024 x 768 pixels.
9.2.5 Text Sonification

Chapter six was presented to the subjects with a soundscape, that is, with background sounds composed exclusively for this piece of text. A composer matched each instance in the text with auditory icons in MP3 format which were recorded on a CD-ROM. Examples of different sounds in the soundscape were footsteps, people talking in a restaurant etc. The soundscape was played back to the subject while he/she was reading at a constant speed. The possibility to synchronize the recorded audio files to self-paced reading was not implemented.

9.2.6 Comprehension Questionnaire

The subjects were prompted to complete a questionnaire after reading chapter one, three, four and five. The questionnaire consisted of ten four-alternative multiple-choice questions for each chapter. The score was measured as a percentage of correct answers. The questionnaires are reproduced as appendices A, B, C, D.

9.2.7 Workload index

For measuring subjective workload the NASA-TLX (Task Load Index)(Hart & Staveland 1988) was used (see appendix E). NASA-TLX measures Mental, Physical demand, Temporal demand, Performance, Effort and Frustration level. Each scale was presented as a 100-millimetre horizontal line with bipolar descriptors at each end. LOW/HIGH was used for all factors except Performance where
GOOD/POOR was used (the higher the value the poorer the performance). Numerical values from 0-100 were assigned to scale positions during data analysis.

9.2.8 Attitude Inventory

In order to examine the readers attitude to reading with the RSVP technique an inventory consisting of five questions were used (see appendix F). The questions measured the subject’s experience of level of Difficulty, Efficiency, Comprehension, Stimulation, and if the ability to download a novel and use the RSVP format would Facilitate their reading. The questions were formulated in comparison with reading from a book. For each question the subject put a mark on a 10-centimetre horizontal scale with bipolar descriptors WORSE/BETTER. Numerical values from –5 to 5 were assigned to positions during data analysis. The value zero implied equal to reading from a book.

9.3 Experimental Design

A repeated-measurement, within-subject design (Hassmén & Koivula 1996:12) was used. The reason for the within-subject design was the limited number of subjects in this study. From a statistical point of view the starting points for the analysis were the null hypotheses outlined in section 9.3.1. The within-subject factor was Reading paradigm (four levels). For the statistical analysis the Repeated-measurement General Linear Model (GLM) in the SPSS Version 10.0 was used. The significance level was set to 5%, and the significance level of multiple comparisons was Bonferroni adjusted (Hassmén & Koivula 1996:122).

9.3.1 Null Hypotheses

1. There is no difference in reading speed between the self-paced reading paradigms.
2. There is no difference in comprehension between the reading paradigms.
3. There is no difference in reading efficiency between the self-paced reading paradigms.
4. Readers experience the same workload regardless of self-paced reading paradigm.
5. Readers have the same attitude to reading regardless of reading paradigm.

---

7 A between-subjects design would have required a larger number of subjects since large differences can be found between users with regard to performance. Or as pointed out in Egan (1988:543): “Differences among people usually account for much more variability in performance than differences in system designs or differences in training procedures”.

9.3.2 Independent Variables

Four different reading conditions were evaluated in the experiment: Book, self-paced RSVP11 (window width eleven characters), self-paced RSVP25 (window width 25 characters) and Sonified RSVP25 (window width 25 characters) with constant reading speed. The Reader 1.0 prototype was used to evaluate the latter three presentation formats.

9.3.3 Dependent Variables

Dependent objective variables were reading speed, comprehension, and reading efficiency. Reading speed was expressed as words per minute (wpm) and comprehension was measured as percentage of questions answered correctly (0–100%) on the Comprehension questionnaire. Reading efficiency was expressed as words per minute.

Dependent subjective variables were workload and attitude measured with NASA-TLX and the Attitude inventory.

9.3.4 Procedure

The subjects were tested one at a time and each experiment took on average 3.5 hours to complete (including a lunch break). Thus, for each subject all conditions were run in one day. After a short introduction, subjects were questioned about their fiction reading habits and they signed a confidentiality agreement and a video recording agreement.

In the test situation the subject was sitting in an armchair with a table on each side for paper and pens. The subject was filmed by video cameras simultaneously from two angles, except when they filled in the questionnaires. One camera, which zoomed in on the PDA, was placed above the subject and the other camera was placed to the left of the subjects’ view to capture when he/she changed reading speed etc.

The experiment was divided into five sessions which lasted from ten to 30 minutes each. If the subject preferred to take a five to ten-minute break between the sessions he/she did so. Between the second and the third session there was a 45 minute break for lunch. Before each session the subject was informed about the procedure for that particular task. The subject did not receive any feedback regarding his/her reading speed in any of the conditions. A brief description of each session is outlined below.

1. In the first session the subjects had to read chapter one of Strindberg’s “Röda Rummet” from a book. They were told to read as they usually read. They were also informed that they had to fill in a questionnaire, about the content of the chapter, and a NASA-TLX form. They were asked to signal when they started and when they finished reading. After finishing the task the subjects completed the Comprehension questionnaire and the NASA-TLX form.
2. In this session which consisted of two practice trials, the RSVP paradigm was introduced. The subjects were to experiment with different speeds for the two formats, window width 11 characters (henceforth RSVP11) and window width 25 characters (henceforth RSVP25) and the two trials were carried out using the same procedure. In this practice session chapter two and five were used. Initially a low speed level was chosen and was then increased until subjects felt they had reached their maximum speed. Next, they were exposed to a speed much higher than their preferred maximum and the speed level was then decreased until the subjects felt comfortable (Björkman & Ekman 1957). This procedure was undertaken as an attempt to increase the subjects’ reading speed. After this session there was a 45-minute break for lunch.

3. In the third session, in which chapter three was read, the subjects read with either RSVP11 or RSVP25. To minimize learning effects the subjects were divided in two groups; Group 1 used the RSVP11 format when reading chapter three and the RSVP25 format when reading chapter four. Group 2 was exposed to the two RSVP formats in reversed order. Initially, the subjects were free to choose speed level. If they wanted to change level while reading they could push the Stop button and choose a faster or slower speed level from the Speed menu. Every speed level change had to be noted on a form. After completion of this task they were asked to fill out the Comprehension questionnaire, the NASA-TLX form and the Attitude questionnaire.

4. In this session chapter four was read. This session was performed with the same procedure as the previous one but with the difference that the other window width was used.

5. In the fifth session the sonified chapter 6 was used. In this task the speed level was set by the experimenter to speed level four (equal to 156 wpm) and the window width was 25 characters.

Thus, the chapters were not read in strict consecutive order since chapter five was used in the second session (the practice session). The reason for this was that a fourth RSVP condition, employing linguistically based segmentation, was included in the experiment from the beginning. This condition was later excluded from the experimental design. An overview of the experiment is listed in table 5.

10 Results

The results of this experiment have been analyzed with regard to reading speed, comprehension, subjective workload and attitude. Due to technical reasons, one of the subjects did not complete the third chapter in the RSVP11 condition. In this
case the reading speed was calculated on the amount of text read (75%) and the comprehension score was extrapolated.

10.1 Objective Findings

No significant differences between the self-paced conditions (Book, RSVP11, and RSVP25) with regard to reading speed were found; neither were differences in comprehension scores significant. Thus, the null hypothesis was kept regarding these measures.

10.1.1 Reading Speed

The number of changes of speed level for the self-paced RSVP conditions varied between zero and eleven for all ten subjects.

The subjects reading speed when reading on paper medium varied between 156 and 402 wpm with an average of 282 wpm. The actual reading time was computed by subtracting the pauses from the total time spent on completing the chapter if applicable (one of the subjects read using the same speed level throughout session four). The time spent on pauses was computed by measuring the time for two speed level changes for each of the ten subjects, and calculating the average of
these 20 pauses. Reading speed for the self-paced conditions are outlined in table 6. Although reading speed was slightly higher for the self-paced RSVP conditions compared to the Book condition the difference was not significant.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Min</th>
<th>Max</th>
<th>Average</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>156</td>
<td>402</td>
<td>282</td>
<td>290</td>
</tr>
<tr>
<td>RSVP11 excl. pause</td>
<td>178</td>
<td>529</td>
<td>363</td>
<td>348</td>
</tr>
<tr>
<td>RSVP25 excl. pause</td>
<td>273</td>
<td>401</td>
<td>317</td>
<td>297</td>
</tr>
</tbody>
</table>

Table 6: Minimum, maximum, mean and median values for reading speed (wpm) for the self-paced conditions.

Figure 4: Box plot presentation of median, 25%–75% percentile and min-max values for reading speed for the self-paced conditions Book, RSVP11 and RSVP25.
10.1.2 Comprehension

Comprehension was calculated as a percentage of correctly answered questions out of 10 (0–100%). There were no significant differences between the reading conditions, however the average score was lower for self-paced RSVP11 (68.5%) compared to the other formats (84–86%).

10.1.3 Reading Efficiency

Reading efficiency was computed as Reading speed times Comprehension for each condition (Jackson & McClelland 1979) and the null hypothesis was kept in part. The difference in Reading Efficiency between the four conditions was significant (F[3,27]=9.88, p=0.03). Pair-wise comparisons showed that a significant difference occurred for Sonified RSVP25 compared to each of the self-paced conditions (p≤0.031). Average and median values for reading efficiency are outlined in table 7. Median, 25%-75% percentile and min-max values for all four conditions are shown in figure 5 on page 28.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>237</td>
<td>290</td>
</tr>
<tr>
<td>RSVP11 excl. pause</td>
<td>249</td>
<td>349</td>
</tr>
<tr>
<td>RSVP25 excl. pause</td>
<td>333</td>
<td>297</td>
</tr>
</tbody>
</table>

Table 7: Average and median values for reading efficiency (wpm) for the self-paced conditions.

Plotting comprehension (%) as a function of reading speed (wpm) for the three self-paced conditions showed a significant negative correlation (R=–0.452, p=0.012), shown in figure 6. Comprehension decreased by 9% as a function of each 100-wpm-increase in speed.

10.2 Subjective Findings

10.2.1 Workload

The null hypothesis, regarding no difference in workload between the three self-paced reading conditions, was partly rejected. All main factors except Physical demand were significant across the three conditions (F[2,18]=6.61, p≤0.007). Pair-wise comparisons between conditions for each factor showed that Mental demand was rated significantly higher for RSVP11 than for Book (p=0.007). Temporal demand was rated significantly higher for the two self-paced RSVP formats than for the Book condition (p≤0.01). Subjects also felt they did not accomplish the goals when reading in RSVP11 condition and rated Performance significantly lower (i.e. significantly worse) than Book (p=0.037). The subjects thought that they had to work mentally and physically harder when reading in both RSVP conditions and
Figure 5: Box plot presentation of median, 25%-75% percentile and min-max values for reading efficiency (wpm) for all four conditions.
Figure 6: Plotting comprehension (%) as a function of reading speed (wpm) for the three self-paced conditions showed a significant negative correlation (R=–0.452, p=0.012).

rated *Effort* significantly higher than Book (p≤0.012). They also rated *Frustration* level significantly higher for both RSVP conditions compared to Book (p≤0.012). Average ratings for the three conditions are shown in figure 7.

### 10.2.2 Attitude

The reader attitude was tested in all four conditions and the null hypothesis was partly rejected. Except for *Efficiency* and *Facilitate* the main factor Reading paradigm was significant (F[3,27]=7.55, p≤0.01). Respecting *Difficulty* pairwise comparison showed that subjects found RSVP11 and RSVP25 significantly more difficult than Book (p≤ 0.003) and RSVP11 significantly more difficult than Sonified RSVP (p≤0.047). Subjective *Comprehension* was significant (p≤0.001) and the subjects
ratted both the self-paced RSVP conditions significantly higher (i.e. significantly worse) than Book \( (p \leq 0.029) \). *Stimulation* when reading in both self-paced conditions of RSVP was considered worse compared to the Book condition \( (p \leq 0.031) \).

### 11 Discussion

Some of the disadvantages of reading on screen according to Shneiderman (see section 2.3) were avoided in this study. Emitted light, flicker and curved display surface are absent on LCDs and there was no need for page turning. Neither uncomfortable reading distance nor fatiguing body posture were present. Furthermore, the text was left justified. On the other hand, the LCD screens have lower resolution than the CRTs.

Reading speed for RSVP on a small screen was found to be slightly, though not significantly, faster than reading speed on paper. This in contrast to earlier studies.

---

**Figure 7:** Ratings (in millimetre) of workload for condition Book, RSVP11, and RSVP25 using the NASA-TLX workload index. Low ratings indicate low workload and good performance.
that compared paper reading to screen reading with regard to reading speed which showed results disfavoring reading on screen (Shneiderman 1998; Edeborg 1999).

Static text presentation on hand-held devices is not optimal due to the limited screen space. This presentation format on a PDA was not included as a condition in the experimental design. Static presentation on a small screen implies more frequent paging or scrolling compared to large screens (Shneiderman 1998). This fact may presumably imply a decrease in reading efficiency for static text presentations on small screens compared to RSVP and book.

Despite their unfamiliarity with RSVP, most of the subjects read faster in the self-paced RSVP conditions than in the Book condition. One explanation for this might be the limitation of eye movements. In addition, the text was aligned to the left which meant that time needed for planning for the next gaze was eliminated. Another consequence of RSVP is that no time is spent on regressions and return sweeps since these are prevented.

One factor that might have reduced reading speed is the text difficulty. The texts were classified as medium to difficult, according to lix, and as the novel was written in the late nineteenth century it contained words which are infrequent in modern Swedish. Subjects may have chosen a lower reading speed due to this since Just and Carpenter (1980) reported longer gaze duration on infrequent words. Another factor that may have prevented subjects from selecting a higher reading speed was a low resolution of the speed level scale (30-40 wpm/speed level unit). A finer grading might have resulted in a faster reading rate.

In session two (practice session) reading speed levels were set using an ascending-descending approach (Björkman & Ekman 1957) and the aim of this procedure was to prevent subjects from subvocalization. In the following sessions the experimenter stressed that a comprehension test would be administered after completion of the chapter. This may have prevented subjects from reading at rapid rates. Almost all subjects chose a lower speed than they could master in session three and four, which may indicate that the general habit of subvocalizing was deeply rooted.

As mentioned before, the term comprehension used in this study is a bit misleading and as Gunnarsson (1982:43) pointed out: “The recall of details indicates that the text has been read but not necessarily that it has been comprehended.”8 The level of learning due to Bloom’s taxonomy was Knowledge since the questionnaires measure if subjects were able to recognize accurate statements.

Reading efficiency was found to be significantly lower for the Sonified RSVP25 condition compared to Book and the self-paced conditions. There were two new factors in the Sonified RSVP25 condition: the soundscape and the constant presentation rate. Accordingly, it is not sure that the low reading efficiency is due to the low speed level.

Different text types imply different reading modes and fiction is most often read in a consecutive order, that is, from the first page to the last. This may not be the case for text when reading goals are other than to fully comprehend every detail

---

8Transl. K.S.
in the text, for instance when reading technical literature. The RSVP technique is, then, probably most suitable for a “linear” reading mode.

The average reading speed on paper was 282 wpm which is in Kump (1999) classified as “average college level”. The amount of fictional books read per year by the participants in this study was lower than expected for readers with an expressed interest in fiction. This low figure may be due to chance, people do not always have the amount of time they would like for reading, or it may indicate that these subjects were not as interested in fiction as desired in the experiment.

A drawback for RSVP is the increase in mental workload. Ratings for most of the factors in the workload inventory were significant in disfavoring RSVP. Another drawback for RSVP was the attitude of the users since they found it less stimulating and more difficult than reading on paper. None of the attitude questions reflected that the multimedia experience was preferable to conventional reading. In the Sonified RSVP25 condition the text presentation rate was set to 156 wpm. This low speed might have been boring since all subjects read at a faster rate in almost all self-paced conditions (one subject read at 156 wpm in Book condition). One explanation for the subjects’ dissatisfaction with RSVP may be their unfamiliarity with RSVP as well as Sonified RSVP, and that the time spent on practicing reading in these formats was very short compared to the time spent on conventional reading. For a fair comparison between the four conditions the subjects would have to spend hundreds of hours of practicing RSVP and Sonified RSVP reading. Another explanation might be the long experiment time. A couple of the subjects found the experiment to be a strenuous exercise and the results might have been different if the experiment had been spread over several sessions run on different days.

The subjects’ unfamiliarity with the RSVP technique may also be a reason for the discrepancy between objective and subjective performance. Objective performance, measured as reading efficiency, was equally good for the Book condition and the self-paced RSVP conditions. However, subjective ratings showed significantly poorer performance for RSVP11 compared to Book.

Another discrepancy of interest is between subjective and objective comprehension. Subjects performed equally well regarding score on the comprehension test in all conditions. However, the experienced comprehension was rated significantly worse for the self-paced RSVP conditions than for Book.

In this study the aim was to maintain a genuine reading situation which was the reason for using fictional texts written by a “real” author. The chapters were read in consecutive order, except for chapter five, which implies that the length and the the difficulty of each chapter varied which may have had an influence on the experiment. The length varied between 3971 and 6309 words for chapters used in the reading sessions (Session 1, 3, 4 and 5). Lix varied between 54 (Book condition) and 42 (RSVP25 and Sonified RSVP25). Neither text difficulty nor text length was constant across experimental conditions. Thus complete experimental control was sacrificed for reality. It is possible that reading speed and comprehension are affected when difficulty and text length are lowered. Nothing has been found in literature regarding changes in reading speed/comprehension that occur as a function
of text difficulty.

Small displays require other presentation formats different from the static text presentation due to the limited screen space. The RSVP technique may be a solution when reading fiction and other text types which are read from the first word to the last since reading speed and comprehension are not affected in a negative way.
Part IV
Concluding Remarks

12 Further Work

Further studies on RSVP using Swedish texts should investigate more sophisticated segmentation, for instance, syntactic segmentation or segmentation into meaningful, coherent units (Pynte & Noizet 1980; Cocklin et al. 1984). This type of segmentation may speed up perceptual processes which in turn may facilitate comprehension and increase reading speed. Furthermore, it would be of interest to examine if an increased window height may facilitate comprehension. That is, to present more than one line at a time and thus provide an enlarged context without an increase in the number of eye movements. It would also be of interest to explore the RSVP technique using text types other than fictional, for example, newspaper articles or scientific texts.

In a real application a continuous rate of presentation would be preferable. The user should also be able to add bookmarks in the text. In a real reading situation it is sometimes necessary to reread small, or large, amounts of text and also to skip passages. So, it would be useful to be able to move forwards or backwards in the text, according to the tape recorder metaphor. Moreover, the application would benefit from the ability to browse the file system for choosing texts.

13 Summary

In this thesis an implementation and an evaluation of a GUI employing the RSVP technique have been studied. In the evaluation reading with the RSVP technique applied on a PDA was compared to reading on paper medium. The study was carried out between the 2 and 12 May 2000 in a user laboratory. Ten subjects, five female and five male, aged between 26 and 43 participated in the study. Four conditions were examined; reading from a book, and reading from three RSVP formats. The texts used were the first six chapters of Strindberg’s “Röda Rummet”. One chapter was presented with a soundscape played back to the user. This study measured reading speed, comprehension, reading efficiency (speed times comprehension), subjective workload and users’ attitude. Subjects comprehended equally well in all four conditions and read equally fast in the three self-paced conditions. Results showed a discrepancy between subjective and objective comprehension and performance with regard to reading from a book and the RSVP conditions. In other words, subjects experienced a poorer comprehension and performance using RSVP, whereas objective comprehension and performance remained unchanged.
References


Fine, Elisabet M. and Eli Peli 1995. Scrolled and rapid serial visual presentation text are read at similar rates by the visually impaired. *Journal of Optical Society of America*, 12(10).


**Online sources**


A Questionnaire Chapter 1

1 I vilken månad utspelar sig händelsen i kapitel ett?
   - April
   - Maj
   - Juni

2 Var utspelar sig samtalet?
   - Vid Skeppsbron
   - I en trädgård på Mosebacke
   - På en krog vid Södermalmtorg

3 Hur många slag slår kyrkklockorna?
   - Sex
   - Sju
   - Åtta

4 Vad heter personerna som samtalar?
   - Ygberg och Lundell
   - Renhjelm och Falk
   - Struve och Falk

5 Vilket yrkesområde ämnar den yngre av personerna slå sig in på?
   - Litteratör
   - Ämbetsman
   - Målare

6 Var går den yngre av båda personerna efter samtalet?
   - Hästholmen
   - Liljeholmen
   - Skeppsholmen

7 Hur länge blir han sittande?
   - Till nästa morgon
   - I fem minuter
   - Till midnatt

8 Var går han sedan?
   - Till Mosebacketorg
   - Till sin hör
   - Till sin vindsvåning

9 För vilken tidning skrev tidigare den äldre av de båda personerna?
   - Gräkappan
   - Rödluvan
   - Svenska Dagbladet

10 För vilken tidning skriver han numera?
    - Gräkappan
    - Rödluvan
    - Svenska Dagbladet
B Questionnaire Chapter 3

1 Vilken tidpunkt på dagen är det i början av kapitlet?
   – Mellan åtta och nio på morgonen
   – Mellan tre och fyra på eftermiddagen
   – Mellan sex och sju på kvällen

2 Hur var vädret?
   – Mulet
   – Soligt
   – Regnigt

3 Vad gör Falk vid Ankdammen?
   – Kastar sten
   – Kastar pinnar
   – Matar ankorna

4 Vad heter stallet han kommer till så småningom?
   – Lill-Jans
   – Per-Ols
   – Stor-Knuts

5 Hur beskriver de båda samtalande personerna vid drivbänkarna?
   – En är ren och en är smutsig
   – En är mager och en är fet
   – En är ung och en är gammal

6 Hur många personer finns i rummet som Falk kliver in i?
   – Två
   – Tre
   – Fyra

7 Vad får Falk syn på vid träsket där skogen börjar?
   – En smart, ung man
   – En hamnbuse
   – Sin bror

8 Vad jobbar Lundell på för tillfället?
   – En vy över Stockholm
   – En altartavla
   – Ett stilleben

9 Var ska sällskapet träffas senare på kvällen?
   – På Grytan
   – På Tennknappen
   – På Röda Rummet

10 Vad gör Olle då vänne Ygberg har somnat?
   – Skissar
   – Somnar
   – Läser
C Questionnaire Chapter 4

1 Var utspelar sig den första delen av kapitel fyra?
   – I sovrummet
   – I matsalen
   – I drängkammaren

2 Hur gammal är Carl Nicolaus Falks hustru?
   – 22 år
   – 27 år
   – 31 år

3 Vilken tid på dagen utspelar sig den första delen?
   – Klockan sju på morgonen
   – Klockan tio på förmiddagen
   – Klockan tio på kvällen

4 Hur länge har paret varit gifta?
   – Två år
   – Fem år
   – Tio år

5 Varför är Falk arg?
   – För att han ska ha främmande
   – För att inte hustrun klätt på sig
   – För att frukosten inte är dukad

6 Vilken tid på dagen utspelar sig den andra delen av kapitlet?
   – Vid lunchtid
   – På eftermiddagen
   – På aften

7 Vad ber Levin Falk om?
   – Låna honom pengar
   – Lägga upp mat på tallriken
   – Dra upp buteljerna

8 Vad heter tidningen där Levin läst artikeln om ”Kollegiet för utbetalandet av Ämbetsmännens löner”?
   – Rödkappan
   – Gråkappan
   – Folkets Fana

9 Vid vilken tidpunkt går Falks gäster?
   – Vid midnatt
   – Vid två-tiden
   – Vid fyra-tiden

10 Vem städar efter festen?
   – Falks hustru
   – Bodbetjänten
   – Falk själv
D Questionnaire Chapter 6

1 Vad gör Sellen i stugan vid Lill-Jans?
   - Målar
   - Läser
   - Predikar

2 Vem av de närvarande blir tillsagd att gå ut och “stampa”?
   - Olle Montanus
   - Renhjelm
   - Sellen

3 Vad ska köpas för pengarna?
   - Färg, bröd och öl
   - En väst
   - Ett par stövlar

4 Vad tycker Lundell att det fattas i tavlan?
   - Moln
   - Karlar
   - En kvinna-figur

5 Hur länge målar de?
   - Till klockan slår fyra
   - Till klockan slår fem
   - Till klockan slår sju

6 På vilken gata bor Falk?
   - Norrlandsgatan
   - Trädgårdsgatan
   - Grevmanngatan

7 Vad av Falks tillhörigheter pantsätter Falk och Sellen?
   - Överrocken
   - Pipan
   - Guldlockklockan

8 Var går de båda sedan?
   - Till Röda Rummet
   - Till Tennknappen
   - Till Grytan

9 Vilken svensk psalm spelar orkestern?
   - Den blomstertid nu kommer
   - Vår Gud är oss en väldig borg
   - Glad jag städse vill bekänna

10 Vad hade hänt Olle när han gick ut för att panta?
    - Han hade blivit nedslagen
    - Han hade blivit bestulen
    - Han hade blivit tagen av polisen
E Workload inventory

Instructions
Place a mark on each scale below that represents the magnitude of each factor in the task you just performed!

1. **Mental demand**
   How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exact or forgiving?
   
   Low .................................................................................................................. High

2. **Physical demand**
   How much physical activity was required (e.g., pushing, pulling, lifting, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
   
   Low .................................................................................................................. High

3. **Temporal demand**
   How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
   
   Low .................................................................................................................. High

4. **Performance**
   How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
   
   Good ................................................................................................................. Poor

5. **Effort**
   How hard did you have to work (mentally and physically) to accomplish your level or performance?
   
   Low .................................................................................................................. High

6. **Frustration level**
   How insincere, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?
   
   Low .................................................................................................................. High
F Attitude inventory

Instruktioner
Markera på varje skala nedan dina skattningar för den nyss genomförda uppgiften.

1. Hur upplevde du svårighetsgraden att läsa på detta sätt i jämförelse med att läsa ur en vanlig bok?
   Lika
   Svårare ________________ | ________________ Lättare

2. Hur pass effektivt upplevde du att det var att läsa på detta sätt i jämförelse med att läsa ur en vanlig bok?
   Lika
   Mindre ________________ | ________________ Mer

3. Hur pass stor läsförmåga upplever du att du har när du läser på detta sätt i jämförelse med att läsa ur en vanlig bok?
   Lika
   Mindre ________________ | ________________ Större

4. Hur pass stimulerande upplevde du att det var läsa på detta sätt i jämförelse med att läsa ur en vanlig bok?
   Lika
   Mindre ________________ | ________________ Mer

5. Hur upplever du att möjligheten att läda ner en roman och läsa på detta sätt skulle underlätta det läsande i jämförelse med att köpa motsvarande bok?
   Lika
   Svårare ________________ | ________________ Lättare


Rapid Serial Visual Presentation (RSVP) is a new fast-content recognition approach which uses electroencephalography to record brain activity elicited by fast bursts of image data. These brain responses are then subjected to machine classification methods to reveal the expert’s “reflex” response to classify images according to their presence or absence of particular targets. The benefit of this method is that images can be presented at high temporal rates (~10 per second), faster than that required for fully conscious detection, facilitating a high throughput of image (screening) material. In